

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

|  |  |  |
|--|--|--|
| <b>(51) International Patent Classification <sup>7</sup> :</b><br><b>C12N 15/62, C07K 14/295, 16/12, A61K 39/118, G01N 33/53, C12Q 1/68, C12N 5/10</b>   | <b>A2</b>  | <b>(11) International Publication Number:</b> <b>WO 00/32794</b><br><br><b>(43) International Publication Date:</b> 8 June 2000 (08.06.00) |
| <b>(21) International Application Number:</b> PCT/CA99/01147<br><b>(22) International Filing Date:</b> 1 December 1999 (01.12.99)<br><br><b>(30) Priority Data:</b><br>60/110,339 1 December 1998 (01.12.98) US<br>60/110,340 1 December 1998 (01.12.98) US<br>60/110,427 1 December 1998 (01.12.98) US<br>60/110,428 1 December 1998 (01.12.98) US<br>60/110,438 1 December 1998 (01.12.98) US<br><br><b>(71) Applicant (for all designated States except US):</b> CONNAUGHT LABORATORIES LIMITED [CA/CA]; 1755 Steeles Avenue West, Toronto, Ontario M2R 3T4 (CA).<br><br><b>(72) Inventors; and</b><br><b>(75) Inventors/Applicants (for US only):</b> MURDIN, Andrew, D. [CA/CA]; 146 Rhodes Circle, Newmarket, Ontario L3X 1V2 (CA). OOMEN, Raymond, P. [CA/CA]; RR No. 1, Schomberg, Ontario L0G 1T0 (CA). WANG, Joe [CA/CA]; 48 29th Street, Etobicoke, Ontario M8W 3A8 (CA).<br><br><b>(74) Agents:</b> WHEELER, Michael, E. et al.; Smart & Biggar, 900-55 Metcalfe Street, P.O. Box 2999, Station D, Ottawa, Ontario K1P 5Y6 (CA). | <b>(81) Designated States:</b> AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).<br><br><b>Published</b><br><i>Without international search report and to be republished upon receipt of that report.</i> |  |
| <b>(54) Title:</b> CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND USES THEREOF<br><br><b>(57) Abstract</b><br><br>The present invention provides purified and isolated polynucleotide molecules that encode <i>Chlamydia</i> polypeptides which can be used in methods to prevent, treat, and diagnose <i>Chlamydia</i> infection. In one form of the invention, the polynucleotide molecules are selected from DNA that encode polypeptides CPN100634 (SEQ ID Nos: 1, 2 and 11), CPN100635 (SEQ ID Nos: 3, 4, 12 and 13), CPN100638 (SEQ ID Nos: 5, 6 and 14), CPN100639 (SEQ ID Nos: 7, 8 and 15), CPN100708 (SEQ ID Nos: 9, 10 and 16).   |  |  |

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

|    |                          |    |  |    |  |    |                          |
|----|--------------------------|----|--|----|--|----|--------------------------|
| AL | Albania                  | ES | Spain                                    | LS | Lesotho                                      | SI | Slovenia                 |
| AM | Armenia                  | FI | Finland                                  | LT | Lithuania                                    | SK | Slovakia                 |
| AT | Austria                  | FR | France                                   | LU | Luxembourg                                   | SN | Senegal                  |
| AU | Australia                | GA | Gabon                                    | LV | Latvia                                       | SZ | Swaziland                |
| AZ | Azerbaijan               | GB | United Kingdom                           | MC | Monaco                                       | TD | Chad                     |
| BA | Bosnia and Herzegovina   | GE | Georgia                                  | MD | Republic of Moldova                          | TG | Togo                     |
| BB | Barbados                 | GH | Ghana                                    | MG | Madagascar                                   | TJ | Tajikistan               |
| BE | Belgium                  | GN | Guinea                                   | MK | The former Yugoslav<br>Republic of Macedonia | TM | Turkmenistan             |
| BF | Burkina Faso             | GR | Greece                                   | ML | Mali   | TR | Turkey                   |
| BG | Bulgaria                 | HU | Hungary                                  | MN | Mongolia                                     | TT | Trinidad and Tobago      |
| BJ | Benin                    | IE | Ireland                                  | MR | Mauritania                                   | UA | Ukraine                  |
| BR | Brazil                   | IL | Israel                                   | MW | Malawi                                       | UG | Uganda                   |
| BY | Belarus                  | IS | Iceland                                  | MX | Mexico                                       | US | United States of America |
| CA | Canada                   | IT | Italy                                    | NE | Niger  | UZ | Uzbekistan               |
| CF | Central African Republic | JP | Japan                                    | NL | Netherlands                                  | VN | Viet Nam                 |
| CG | Congo                    | KE | Kenya                                    | NO | Norway                                       | YU | Yugoslavia               |
| CH | Switzerland              | KG | Kyrgyzstan                               | NZ | New Zealand                                  | ZW | Zimbabwe                 |
| CI | Côte d'Ivoire            | KP | Democratic People's<br>Republic of Korea | PL | Poland                                       |    |                          |
| CM | Cameroon                 | KR | Republic of Korea                        | PT | Portugal                                     |    |                          |
| CN | China                    | KZ | Kazakstan                                | RO | Romania                                      |    |                          |
| CU | Cuba                     | LC | Saint Lucia                              | RU | Russian Federation                           |    |                          |
| CZ | Czech Republic           | LI | Liechtenstein                            | SD | Sudan  |    |                          |
| DE | Germany                  | LK | Sri Lanka                                | SE | Sweden                                       |    |                          |
| DK | Denmark                  | LR | Liberia                                  | SG | Singapore                                    |    |                          |
| EE | Estonia                  |    |  |    |  |    |                          |

**TITLE OF INVENTION**

CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND  
USES THEREOF

**5 REFERENCE TO RELATED APPLICATIONS**

This application claims the benefit of U.S.  
Provisional Application No. 60/110,427, filed December 1, 1998,  
U.S. Provisional Application No. 60/110,438, filed December 1,  
1998, U.S. Provisional Application No. 60/110,339, filed December  
10 1, 1998, U.S. Provisional Application No. 60/110,428, filed  
December 1, 1998, U.S. Provisional Application No. 60/110,340,  
filed December 1, 1998.

**FIELD OF INVENTION**

15 The present invention relates to *Chlamydia* antigens  
and corresponding DNA molecules, which can be used to prevent  
and treat *Chlamydia* infection in mammals, such as humans.

**BACKGROUND OF THE INVENTION**

20 *Chlamydiae* are prokaryotes. They exhibit morphologic  
and structural similarities to gram-negative bacteria including  
a trilaminar outer membrane, which contains lipopolysaccharide  
and several membrane proteins that are structurally and  
functionally analogous to proteins found in *E coli*. They are  
25 obligate intra-cellular parasites with a unique biphasic life  
cycle consisting of a metabolically inactive but infectious  
extracellular stage and a replicating but non-infectious  
intracellular stage. The replicative stage of the life-cycle  
takes place within a membrane-bound inclusion which sequesters  
30 the bacteria away from the cytoplasm of the infected host cell.

*C. pneumoniae* is a common human pathogen, originally  
described as the TWAR strain of *Chlamydia psittaci* but  
subsequently recognised to be a new species. *C. pneumoniae* is  
antigenically, genetically and morphologically distinct from



other chlamydia species (*C. trachomatis*, *C. pecorum* and *C. psittaci*). It shows 10% or less DNA sequence homology with either of *C. trachomatis* or *C. psittaci*.

*C. pneumoniae* is a common cause of community acquired pneumonia, only less frequent than *Streptococcus pneumoniae* and *Mycoplasma pneumoniae* (Grayston et al. (1995) Journal of Infectious Diseases 168:1231; Campos et al. (1995) Investigation of Ophthalmology and Visual Science 36:1477). It can also cause upper respiratory tract symptoms and disease, including bronchitis and sinusitis (Grayston et al. (1995) Journal of Infectious Diseases 168:1231; Grayston et al (1990) Journal of Infectious Diseases 161:618; Marrie (1993) Clinical Infectious Diseases. 18:501; Wang et al (1986) Chlamydial infections). Cambridge University Press, Cambridge. p. 329 The great majority of the adult population (over 60%) has antibodies to *C. pneumoniae* (Wang et al (1986) Chlamydial infections. Cambridge University Press, Cambridge. p. 329), indicating past infection which was unrecognized or asymptomatic.

*C. pneumoniae* infection usually presents as an acute respiratory disease (i.e., cough, sore throat, hoarseness, and fever; abnormal chest sounds on auscultation). For most patients, the cough persists for 2 to 6 weeks, and recovery is slow. In approximately 10% of these cases, upper respiratory tract infection is followed by bronchitis or pneumonia. Furthermore, during a *C. pneumoniae* epidemic, subsequent co-infection with pneumococcus has been noted in about half of these pneumonia patients, particularly in the infirm and the elderly. As noted above, there is more and more evidence that *C. pneumoniae* infection is also linked to diseases other than respiratory infections.

The reservoir for the organism is presumably people. In contrast to *C. psittaci* infections, there is no known bird or animal reservoir. Transmission has not been clearly defined. It may result from direct contact with secretions, from fomites,

or from airborne spread. There is a long incubation period, which may last for many months. Based on analysis of epidemics, *C. pneumoniae* appears to spread slowly through a population (case-to-case interval averaging 30 days) because infected persons are inefficient transmitters of the organism. Susceptibility to *C. pneumoniae* is universal. Reinfections occur during adulthood, following the primary infection as a child. *C. pneumoniae* appears to be an endemic disease throughout the world, noteworthy for superimposed intervals of increased incidence (epidemics) that persist for 2 to 3 years. *C. trachomatis* infection does not confer cross-immunity to *C. pneumoniae*. Infections are easily treated with oral antibiotics, tetracycline or erythromycin (2 g/d, for at least 10 to 14 d). A recently developed drug, azithromycin, is highly effective as a single-dose therapy against chlamydial infections.

In most instances, *C. pneumoniae* infection is often mild and without complications, and up to 90% of infections are subacute or unrecognized. Among children in industrialized countries, infections have been thought to be rare up to the age of 5 y, although a recent study (E Normann et al, Chlamydia pneumoniae in children with acute respiratory tract infections, Acta Paediatrica, 1998, Vol 87, Iss 1, pp 23-27) has reported that many children in this age group show PCR evidence of infection despite being seronegative, and estimates a prevalence of 17-19% in 2-4 y olds. In developing countries, the seroprevalence of *C. pneumoniae* antibodies among young children is elevated, and there are suspicions that *C. pneumoniae* may be an important cause of acute lower respiratory tract disease and mortality for infants and children in tropical regions of the world.

From seroprevalence studies and studies of local epidemics, the initial *C. pneumoniae* infection usually happens between the ages of 5 and 20 y. In the USA, for example, there

are estimated to be 30,000 cases of childhood pneumonia each year caused by *C. pneumoniae*. Infections may cluster among groups of children or young adults (e.g., school pupils or military conscripts).

5            *C. pneumoniae* causes 10 to 25% of community-acquired lower respiratory tract infections (as reported from Sweden, Italy, Finland, and the USA). During an epidemic, *C. pneumonia* infection may account for 50 to 60% of the cases of pneumonia. During these periods, also, more episodes of mixed infections  
10 with *S. pneumoniae* have been reported.

Reinfection during adulthood is common; the clinical presentation tends to be milder. Based on population seroprevalence studies, there tends to be increased exposure with age, which is particularly evident among men. Some  
15 investigators have speculated that a persistent, asymptomatic *C. pneumoniae* infection state is common.

In adults of middle age or older, *C. pneumoniae* infection may progress to chronic bronchitis and sinusitis. A study in the USA revealed that the incidence of pneumonia caused  
20 by *C. pneumoniae* in persons younger than 60 years is 1 case per 1,000 persons per year; but in the elderly, the disease incidence rose three-fold. *C. pneumoniae* infection rarely leads to hospitalization, except in patients with an underlying illness.

25            Of considerable importance is the association of atherosclerosis and *C. pneumoniae* infection. There are several epidemiological studies showing a correlation of previous infections with *C. pneumoniae* and heart attacks, coronary artery and carotid artery disease (Saikku et al. (1988) Lancet; ii:983;  
30 Thom et al. (1992) JAMA 268:68; Linnanmaki et al. (1993), Circulation 87:1030; Saikku et al. (1992) Annals Internal Medicine 116:273; Melnick et al (1993) American Journal of Medicine 95:499). Moreover, the organisms has been detected in atheromas and fatty streaks of the coronary, carotid, peripheral

arteries and aorta (Shor et al. (1992) South African. Medical Journal 82:158; Kuo et al. (1993) Journal of Infectious Diseases 167:841; Kuo et al. (1993) Arteriosclerosis and Thrombosis 13:1500; Campbell et al (1995) Journal of Infectious Diseases 5 172:585; Chiu et al. Circulation, 1997 (In Press)). Viable *C. pneumoniae* has been recovered from the coronary and carotid artery (Ramirez et al (1996) Annals of Internal Medicine 125:979; Jackson et al. Abst. K121, p272, 36<sup>th</sup> ICAAC, 15-18 Sept. 1996, New Orleans). Furthermore, it has been shown that 10 *C. pneumoniae* can induce changes of atherosclerosis in a rabbit model (Fong et al (1997) Journal of Clinical Microbiology 35:48). Taken together, these results indicate that it is highly probable that *C. pneumoniae* can cause atherosclerosis in humans, though the epidemiological importance of chlamydial 15 atherosclerosis remains to be demonstrated.

A number of recent studies have also indicated an association between *C. pneumoniae* infection and asthma. Infection has been linked to wheezing, asthmatic bronchitis, adult-onset asthma and acute exacerbations of asthma in adults, 20 and small-scale studies have shown that prolonged antibiotic treatment was effective at greatly reducing the severity of the disease in some individuals (Hahn DL, et al. Evidence for *Chlamydia pneumoniae* infection in steroid-dependent asthma. Ann Allergy Asthma Immunol. 1998 Jan; 80(1): 45-49.; Hahn DL, et 25 al. Association of *Chlamydia pneumoniae* IgA antibodies with recently symptomatic asthma. Epidemiol Infect. 1996 Dec; 117(3): 513-517; Bjornsson E, et al. Serology of chlamydia in relation to asthma and bronchial hyperresponsiveness. Scand J Infect Dis. 1996; 28(1): 63-69.; Hahn DL. Treatment of *Chlamydia* 30 *pneumoniae* infection in adult asthma: a before-after trial. J Fam Pract. 1995 Oct; 41(4): 345-351.; Allegra L, et al. Acute exacerbations of asthma in adults: role of *Chlamydia pneumoniae* infection. Eur Respir J. 1994 Dec; 7(12): 2165-2168.; Hahn DL, et al. Association of *Chlamydia pneumoniae* (strain TWAR)

infection with wheezing, asthmatic bronchitis, and adult-onset asthma. JAMA. 1991 Jul 10; 266(2): 225-230).

In light of these results a protective vaccine against *C. pneumoniae* infection would be of considerable importance.

5 There is not yet an effective vaccine for any human chlamydial infection. It is conceivable that an effective vaccine can be developed using physically or chemically inactivated Chlamydiae. However, such a vaccine does not have a high margin of safety. In general, safer vaccines are made by genetically manipulating  
10 the organism by attenuation or by recombinant means.

Accordingly, a major obstacle in creating an effective and safe vaccine against human chlamydial infection has been the paucity of genetic information regarding Chlamydia, specifically *C. pneumoniae*.

15 Studies with *C. trachomatis* and *C. psittaci* indicate that safe and effective vaccine against Chlamydia is an attainable goal. For example, mice which have recovered from a lung infection with *C. trachomatis* are protected from infertility induced by a subsequent vaginal challenge (Pal et  
20 al. (1996) Infection and Immunity 64:5341). Similarly, sheep immunized with inactivated *C. psittaci* were protected from subsequent chlamydial-induced abortions and stillbirths (Jones et al. (1995) Vaccine 13:715). Protection from chlamydial infections has been associated with Th1 immune responses,  
25 particularly the induction of INF $\gamma$  - producing CD4+T-cells (Igiertsemes et al. (1993) Immunology 5:317). The adoptive transfer of CD4+ cell lines or clones to nude or SCID mice conferred protection from challenge or cleared chronic disease (Igiertseme et al (1993) Regional Immunology 5:317; Magee et al  
30 (1993) Regional Immunology 5: 305), and *in vivo* depletion of CD4+ T cells exacerbated disease post-challenge (Landers et al (1991) Infection & Immunity 59:3774; Magee et al (1995) Infection & Immunity 63:516). However, the presence of sufficiently high titres of neutralising antibody at mucosal

surfaces can also exert a protective effect (Cotter et al. (1995) Infection and Immunity 63:4704).

Antigenic variation within the species *C. pneumoniae* is not well documented due to insufficient genetic information, though variation is expected to exist based on *C. trachomatis*. Serovars of *C. trachomatis* are defined on the basis of antigenic variation in the major outer membrane protein (MOMP), but published *C. pneumoniae* MOMP gene sequences show no variation between several diverse isolates of the organism (Campbell et al 10 (1990) Infection and Immunity 58:93; McCafferty et al (1995) Infection and Immunity 63:2387-9; Knudsen et al (1996) Third Meeting of the European Society for Chlamydia Research, Vienna). Regions of the protein known to be conserved in other chlamydial MOMPs are conserved in *C. pneumoniae* (Campbell et al (1990) 15 Infection and Immunity 58:93; McCafferty et al (1995) Infection and Immunity 63:2387-9). One study has described a strain of *C. pneumoniae* with a MOMP of greater than usual molecular weight, but the gene for this has not been sequenced (Grayston et al. (1995) Journal of Infectious Diseases 168:1231). Partial 20 sequences of outer membrane protein 2 from nine diverse isolates were also found to be invariant (Ramirez et al (1996) Annals of Internal Medicine 125:979). The genes for HSP60 and HSP70 show little variation from other chlamydial species, as would be expected. The gene encoding a 76kDa antigen has been cloned 25 from a single strain of *C. pneumoniae*. It has no significant similarity with other known chlamydial genes (Marrie (1993) Clinical Infectious Diseases. 18:501).

Many antigens recognised by immune sera to *C. pneumoniae* are conserved across all chlamydiae, but 98kDa, 30 76 kDa and 54 kDa proteins appear to be *C. pneumoniae*-specific (Campos et al. (1995) Investigation of Ophthalmology and Visual Science 36:1477; Marrie (1993) Clinical Infectious Diseases. 18:501; Wiedmann-Al-Ahmad M, et al. Reactions of polyclonal and neutralizing anti-p54 monoclonal antibodies with an isolated,

species-specific 54-kilodalton protein of *Chlamydia pneumoniae*.  
Clin Diagn Lab Immunol. 1997 Nov; 4(6): 700-704).

Immunoblotting of isolates with sera from patients does show  
variation of blotting patterns between isolates, indicating that  
5 serotypes *C. pneumoniae* may exist (Grayston et al. (1995)  
Journal of Infectious Diseases 168:1231; Ramirez et al (1996)  
Annals of Internal Medicine 125:979). However, the results are  
potentially confounded by the infection status of the patients,  
since immunoblot profiles of a patient's sera change with time  
10 post-infection. An assessment of the number and relative  
frequency of any serotypes, and the defining antigens, is not  
yet possible.

Accordingly, a need exists for identifying and  
isolating polynucleotide sequences of *C. pneumoniae* for use in  
15 preventing and treating *Chlamydia* infection.

#### SUMMARY OF THE INVENTION

The present invention provides purified and isolated  
polynucleotide molecules that encode *Chlamydia* polypeptides  
20 which can be used in methods to prevent, treat, and diagnose  
*Chlamydia* infection. In one form of the invention, the  
polynucleotide molecules are selected from DNA that encode  
polypeptides CPN100634 (SEQ ID Nos: 1 and 2), CPN100635 (SEQ ID  
Nos: 3 and 4), CPN100638 (SEQ ID Nos: 5 and 6), CPN100639 (SEQ  
25 ID Nos: 7 and 8), and CPN100708 (SEQ ID Nos: 9 and 10).

Another form of the invention provides polypeptides  
corresponding to the isolated DNA molecules. The amino acid  
sequences of the corresponding encoded polypeptides are shown  
for CPN100634 as SEQ ID No: 11, CPN100635 as SEQ ID Nos: 12 and  
30 13, CPN100638 as SEQ ID No: 14, CPN100639 as SEQ ID No: 15, and  
CPN100708 as SEQ ID No: 16.

Those skilled in the art will readily understand that the  
invention, having provided the polynucleotide sequences encoding  
*Chlamydia* polypeptides, also provides polynucleotides encoding

fragments derived from such peptides. Moreover, the invention is understood to provide mutants and derivatives of such polypeptides and fragments derived therefrom, which result from the addition, deletion, or substitution of non-essential amino acids as described herein. Those skilled in the art would also readily understand that the invention, having provided the polynucleotide sequences encoding *Chlamydia* polypeptides, further provides monospecific antibodies that specifically bind to such polypeptides.

10       The present invention has wide application and includes expression cassettes, vectors, and cells transformed or transfected with the polynucleotides of the invention. Accordingly, the present invention further provides (i) a method for producing a polypeptide of the invention in a recombinant host system and related expression cassettes, vectors, and transformed or transfected cells; (ii) a vaccine, or a live vaccine vector such as a pox virus, *Salmonella typhimurium*, or *Vibrio cholerae* vector, containing a polynucleotide of the invention, such vaccines and vaccine vectors being useful for, 20 e.g., preventing and treating *Chlamydia* infection, in combination with a diluent or carrier, and related pharmaceutical compositions and associated therapeutic and/or prophylactic methods; (iii) a therapeutic and/or prophylactic use of an RNA or DNA molecule of the invention, either in a naked form or formulated with a delivery vehicle, a polypeptide or combination of polypeptides, or a monospecific antibody of the invention, and related pharmaceutical compositions; (iv) a method for diagnosing the presence of *Chlamydia* in a biological sample, which can involve the use of a DNA or RNA molecule, a 25 monospecific antibody, or a polypeptide of the invention; and (v) a method for purifying a polypeptide of the invention by antibody-based affinity chromatography.



**BRIEF DESCRIPTION OF THE DRAWINGS**

The present invention will be further understood from the following description with reference to the drawings, in which:

Figure 1 shows the nucleotide sequence of the CPN100634 (SEQ ID No: 1 - entire sequence and SEQ ID No: 2 - coding sequence) and the deduced amino acid sequence of the CPN100634 protein from *Chlamydia pneumoniae* (SEQ ID No: 11).

Figure 2 shows the restriction enzyme analysis of the gene encoding the *C. pneumoniae* CPN100634 gene.

10 Figure 3 shows the nucleotide sequence of the CPN100635 (SEQ ID No: 3 - entire sequence and SEQ ID No: 4 - coding sequence) and the deduced amino acid sequence of the CPN100635 protein from *Chlamydia pneumoniae* (SEQ ID No: 12 - entire amino acid sequence corresponding to the open reading frame, and 13 -  
15 the post-translationally processed polypeptide).

Figure 4 shows the restriction enzyme analysis of the gene encoding the *C. pneumoniae* CPN100635 gene.

Figure 5 shows the nucleotide sequence of the CPN100638 (SEQ ID No: 5 - entire sequence and SEQ ID No: 6 - coding  
20 sequence) and the deduced amino acid sequence of the CPN100638 protein from *Chlamydia pneumoniae* (SEQ ID No: 14). The sequence is encoded on the negative strand.

Figure 6 shows the restriction enzyme analysis of the gene encoding the *C. pneumoniae* CPN100638 gene.

25 Figure 7 shows the nucleotide sequence of the CPN100639 (SEQ ID No: 7 - entire sequence and SEQ ID No: 8 - coding sequence) and the deduced amino acid sequence of the CPN100639 protein from *Chlamydia pneumoniae* (SEQ ID No: 15).

Figure 8 shows the restriction enzyme analysis of the  
30 gene encoding the *C. pneumoniae* CPN100639 gene.

Figure 9 shows the nucleotide sequence of the CPN100708 (SEQ ID No: 9 - entire sequence and SEQ ID No: 10 - coding sequence coded for on the negative strand) and the deduced amino

acid sequence of the CPN100708 protein from *Chlamydia pneumoniae* (SEQ ID No: 16).

Figure 10 shows the restriction enzyme analysis of the gene encoding the *C. pneumoniae* CPN100708 gene.

5        Figures 11 through 15 show an identification of T and B cell epitopes from the amino acid sequences shown in the foregoing figures.

#### DETAILED DESCRIPTION OF INVENTION

10        Open reading frames (ORFs) encoding chlamydial polypeptides have been identified from the *C. pneumoniae* genome. These polypeptides include polypeptides found permanently in the bacterial membrane structure, polypeptides present in the external vicinity of the bacterial membrane, polypeptides found  
15 permanently in the inclusion membrane structure, polypeptides present in the external vicinity of the inclusion membrane, and polypeptides released into the cytoplasm of the infected cell. These polypeptides can be used to prevent and treat *Chlamydia* infection.

20        According to a first aspect of the invention, isolated polynucleotides are provided which encode the precursor and mature forms of *Chlamydia* polypeptides, whose amino acid sequences are selected from the group consisting of: SEQ ID Nos: 11 to 16.

25        The term "isolated polynucleotide" is defined as a polynucleotide removed from the environment in which it naturally occurs. For example, a naturally-occurring DNA molecule present in the genome of a living bacteria or as part of a gene bank is not isolated, but the same molecule separated  
30 from the remaining part of the bacterial genome, as a result of, e.g., a cloning event (amplification), is isolated. Typically, an isolated DNA molecule is free from DNA regions (e.g., coding regions) with which it is immediately contiguous at the 5' or 3' end, in the naturally occurring genome. Such isolated

polynucleotides may be part of a vector or a composition and still be defined as isolated in that such a vector or composition is not part of the natural environment of such polynucleotide.

- 5       The polynucleotide of the invention is either RNA or DNA (cDNA, genomic DNA, or synthetic DNA), or modifications, variants, homologs or fragments thereof. The DNA is either double-stranded or single-stranded, and, if single-stranded, is either the coding strand or the non-coding (anti-sense) strand.
- 10 Any one of the sequences that encode the polypeptides of the invention as shown in SEQ ID Nos: 1 to 10 is (a) a coding sequence, (b) a ribonucleotide sequence derived from transcription of (a), or (c) a coding sequence which uses the redundancy or degeneracy of the genetic code to encode the same
- 15 polypeptides. By "polypeptide" or "protein" is meant any chain of amino acids, regardless of length or post-translational modification (e.g., glycosylation or phosphorylation). Both terms are used interchangeably in the present application.

Consistent with the first aspect of the invention, amino

20 acid sequences are provided which are homologous to any one of SEQ ID Nos: 11 to 16. As used herein, "homologous amino acid sequence" is any polypeptide which is encoded, in whole or in part, by a nucleic acid sequence which hybridizes at 25-35°C below critical melting temperature ( $T_m$ ), to any portion of the

25 nucleic acid sequences of SEQ ID Nos: 1 to 10. A homologous amino acid sequence is one that differs from an amino acid sequence shown in any one of SEQ ID Nos: 11 to 16 by one or more conservative amino acid substitutions. Such a sequence also encompass serotypic variants (defined below) as well as

30 sequences containing deletions or insertions which retain inherent characteristics of the polypeptide such as immunogenicity. Preferably, such a sequence is at least 75%, more preferably 80%, and most preferably 90% identical to any one of SEQ ID Nos: 11 to 16.

Homologous amino acid sequences include sequences that are identical or substantially identical to SEQ ID Nos: 11 to 16. By "amino acid sequence substantially identical" is meant a sequence that is at least 90%, preferably 95%, more preferably 5 97%, and most preferably 99% identical to an amino acid sequence of reference and that preferably differs from the sequence of reference by a majority of conservative amino acid substitutions.

Conservative amino acid substitutions are substitutions 10 among amino acids of the same class. These classes include, for example, amino acids having uncharged polar side chains, such as asparagine, glutamine, serine, threonine, and tyrosine; amino acids having basic side chains, such as lysine, arginine, and histidine; amino acids having acidic side chains, such as 15 aspartic acid and glutamic acid; and amino acids having nonpolar side chains, such as glycine, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan, and cysteine.

Homology is measured using sequence analysis software 20 such as Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, WI 53705. Amino acid sequences are aligned to maximize identity. Gaps may be artificially introduced into the sequence to attain proper alignment. Once 25 the optimal alignment has been set up, the degree of homology is established by recording all of the positions in which the amino acids of both sequences are identical, relative to the total number of positions.

Homologous polynucleotide sequences are defined in a 30 similar way. Preferably, a homologous sequence is one that is at least 45%, more preferably 60%, and most preferably 85% identical to any one of coding sequences SEQ ID Nos: 1 to 10.

Consistent with the first aspect of the invention, polypeptides having a sequence homologous to any one of SEQ ID

Nos: 11 to 16 include naturally-occurring allelic variants, as well as mutants or any other non-naturally occurring variants that retain the inherent characteristics of the polypeptide of SEQ ID Nos: 11 to 16.

5           As is known in the art, an allelic variant is an alternate form of a polypeptide that is characterized as having a substitution, deletion, or addition of one or more amino acids that does not alter the biological function of the polypeptide. By "biological function" is meant the function of the  
10 polypeptide in the cells in which it naturally occurs, even if the function is not necessary for the growth or survival of the cells. For example, the biological function of a porin is to allow the entry into cells of compounds present in the extracellular medium. Biological function is distinct from  
15 antigenic property. A polypeptide can have more than one biological function.

          Allelic variants are very common in nature. For example, a bacterial species such as *C. pneumoniae*, is usually represented by a variety of strains that differ from each other  
20 by minor allelic variations. Indeed, a polypeptide that fulfills the same biological function in different strains can have an amino acid sequence (and polynucleotide sequence) that is not identical in each of the strains. Despite this variation, an immune response directed generally against many  
25 allelic variants has been demonstrated. In studies of the *Chlamydial* MOMP antigen, cross-strain antibody binding plus neutralization of infectivity occurs despite amino acid sequence variation of MOMP from strain to strain, indicating that the MOMP, when used as an immunogen, is tolerant of amino acid  
30 variations.

          Polynucleotides encoding homologous polypeptides or allelic variants are retrieved by polymerase chain reaction (PCR) amplification of genomic bacterial DNA extracted by conventional methods. This involves the use of synthetic

oligonucleotide primers matching upstream and downstream of the 5' and 3' ends of the encoding domain. Suitable primers are designed according to the nucleotide sequence information provided in SEQ ID Nos: 1 to 10. The procedure is as follows: a primer is selected which consists of 10 to 40, preferably 15 to 25 nucleotides. It is advantageous to select primers containing C and G nucleotides in a proportion sufficient to ensure efficient hybridization; *i.e.*, an amount of C and G nucleotides of at least 40%, preferably 50% of the total nucleotide content. A standard PCR reaction contains typically 0.5 to 5 Units of Taq DNA polymerase per 100  $\mu$ L, 20 to 200  $\mu$ M deoxynucleotide each, preferably at equivalent concentrations, 0.5 to 2.5 MM magnesium over the total deoxynucleotide concentration,  $10^5$  to  $10^6$  target molecules, and about 20 pmol of each primer. About 25 to 50 PCR cycles are performed, with an annealing temperature 15°C to 5°C below the true  $T_m$  of the primers. A more stringent annealing temperature improves discrimination against incorrectly annealed primers and reduces incorporation of incorrect nucleotides at the 3' end of primers. A denaturation temperature of 95°C to 97°C is typical, although higher temperatures may be appropriate for dematuration of G+C-rich targets. The number of cycles performed depends on the starting concentration of target molecules, though typically more than 40 cycles is not recommended as non-specific background products tend to accumulate.

An alternative method for retrieving polynucleotides encoding homologous polypeptides or allelic variants is by hybridization screening of a DNA or RNA library. Hybridization procedures are well-known in the art and are described in Ausubel *et al.*, (Ausubel *et al.*, Current Protocols in Molecular Biology, John Wiley & Sons Inc., 1994), Silhavy *et al.* (Silhavy *et al.* Experiments with Gene Fusions, Cold Spring Harbor Laboratory Press, 1984), and Davis *et al.* (Davis *et al.* A Manual

for Genetic Engineering: Advanced Bacterial Genetics, Cold Spring Harbor Laboratory Press, 1980)). Important parameters for optimizing hybridization conditions are reflected in a formula used to obtain the critical melting temperature above which two complementary DNA strands separate from each other (Casey & Davidson, Nucl. Acid Res. (1977) 4:1539). For polynucleotides of about 600 nucleotides or larger, this formula is as follows:  $T_m = 81.5 + 0.5 \times (\% \text{ G+C}) + 1.6 \log (\text{positive ion concentration}) - 0.6 \times (\% \text{ formamide})$ . Under appropriate stringency conditions, hybridization temperature ( $T_h$ ) is approximately 20 to 40°C, 20 to 25°C, or, preferably 30 to 40°C below the calculated  $T_m$ . Those skilled in the art will understand that optimal temperature and salt conditions can be readily determined.

For the polynucleotides of the invention, stringent conditions are achieved for both pre-hybridizing and hybridizing incubations (i) within 4-16 hours at 42°C, in 6 x SSC containing 50% formamide, or (ii) within 4-16 hours at 65°C in an aqueous 6 x SSC solution (1 M NaCl, 0.1 M sodium citrate (pH 7.0)).

Useful homologs and fragments thereof that do not occur naturally are designed using known methods for identifying regions of an antigen that are likely to tolerate amino acid sequence changes and/or deletions. As an example, homologous polypeptides from different species are compared; conserved sequences are identified. The more divergent sequences are the most likely to tolerate sequence changes. Homology among sequences may be analyzed using the BLAST homology searching algorithm of Altschul et al., Nucleic Acids Res. 25:3389-3402 (1997). Alternatively, sequences are modified such that they become more reactive to T- and/or B-cells. (See Figures 11 to 15 below for identification of T- and B- epitopes). Yet another alternative is to mutate a particular amino acid residue or sequence within the polypeptide *in vitro*, then screen the mutant

polypeptides for their ability to prevent or treat Chlamydia infection according to the method outlined below.

A person skilled in the art will readily understand that by following the screening process of this invention, it will be determined without undue experimentation whether a particular homolog of any of SEQ ID Nos: 11 to 16 may be useful in the prevention or treatment of Chlamydia infection. The screening procedure comprises the steps:

- (i) immunizing an animal, preferably mouse, with the test homolog or fragment;
- (ii) inoculating the immunized animal with Chlamydia; and,
- (iii) selecting those homologs or fragments which confer protection against Chlamydia.

By "conferring protection" is meant that there is a reduction in severity of any of the effects of Chlamydia infection, in comparison with a control animal which was not immunized with the test homolog or fragment.

It has been previously demonstrated (Yang, Z. P., Chi, E. Y., Kuo, C. C. and Grayston, J. T. 1993. A mouse model of *C. pneumoniae* strain TWAR pneumonitis. 61(5):2037-2040) that mice are susceptible to intranasal infection with different isolates of *C. pneumoniae*. Strain AR-39 (Chi, E. Y., Kuo, C. C. and Grayston, J. T. , 1987. Unique ultrastructure in the elementary body of Chlamydia sp. strain TWAR. J. Bacteriol. 169(8):3757-63) was used in Balb/c mice as a challenge infection model to examine the capacity of chlamydia gene products delivered as naked DNA to elicit a protective response against a sublethal *C. pneumoniae* lung infection. Protective immunity is defined as an accelerated clearance of pulmonary infection.

Groups of 7 to 9 week old male Balb/c mice (6 to 10 per group) were immunized intramuscularly (i.m.) plus intranasally (i.n.) with plasmid DNA containing the coding sequence of a *C. pneumoniae* polypeptide. Saline or the plasmid vector lacking



an inserted chlamydial gene was given to groups of control animals.

For i.m. immunization alternate left and right quadriceps were injected with 100µg of DNA in 50µl of PBS on three 5 occasions at 0, 3 and 6 weeks. For i.n. immunization, anaesthetized mice aspirated 50µl of PBS containing 50 µg DNA on three occasions at 0, 3 and 6 weeks. At week 8, immunized mice were inoculated i.n. with  $5 \times 10^5$  IFU of *C. pneumoniae*, strain AR39 in 100µl of SPG buffer to test their ability to limit the 10 growth of a sublethal *C. pneumoniae* challenge.

Lungs were taken from mice at day 9 post-challenge and immediately homogenised in SPG buffer (7.5% sucrose, 5mM glutamate, 12.5mM phosphate pH7.5). The homogenate was stored frozen at -70°C until assay. Dilutions of the homogenate were 15 assayed for the presence of infectious chlamydia by inoculation onto monolayers of susceptible cells. The inoculum was centrifuged onto the cells at 3000rpm for 1 hour, then the cells were incubated for three days at 35°C in the presence of 1µg/ml cycloheximide. After incubation the monolayers were fixed with 20 formalin and methanol then immunoperoxidase stained for the presence of chlamydial inclusions using convalescent sera from rabbits infected with *C. pneumoniae* and metal-enhanced DAB as a peroxidase substrate.

Consistent with the first aspect of the invention, 25 polypeptide derivatives are provided that are partial sequences of SEQ ID Nos: 11 to 16, partial sequences of polypeptide sequences homologous to SEQ ID Nos: 11 to 16, polypeptides derived from full-length polypeptides by internal deletion, and fusion proteins.

30 It is an accepted practice in the field of immunology to use fragments and variants of protein immunogens as vaccines, as all that is required to induce an immune response to a protein is a small (e.g., 8 to 10 amino acid) immunogenic region of the

protein. Various short synthetic peptides corresponding to surface-exposed antigens of pathogens other than *Chlamydia* have been shown to be effective vaccine antigens against their respective pathogens, e.g. an 11 residue peptide of murine mammary tumor virus (Casey & Davidson, Nucl. Acid Res. (1977) 4:1539), a 16-residue peptide of Semliki Forest virus (Snijders et al., 1991. J. Gen. Virol. 72:557-565), and two overlapping peptides of 15 residues each from canine parvovirus (Langeveld et al., Vaccine 12(15):1473-1480, 1994).

10 Accordingly, it will be readily apparent to one skilled in the art, having read the present description, that partial sequences of SEQ ID Nos: 11 to 16 or their homologous amino acid sequences are inherent to the full-length sequences and are taught by the present invention. Such polypeptide fragments 15 preferably are at least 12 amino acids in length.

Advantageously, they are at least 20 amino acids, preferably at least 50 amino acids, more preferably at least 75 amino acids, and most preferably at least 100 amino acids in length.

Polynucleotides of 30 to 600 nucleotides encoding partial 20 sequences of sequences homologous to SEQ ID Nos: 11 to 16 are retrieved by PCR amplification using the parameters outlined above and using primers matching the sequences upstream and downstream of the 5' and 3' ends of the fragment to be amplified. The template polynucleotide for such amplification 25 is either the full length polynucleotide homologous to one of SEQ ID Nos: 1 to 10, or a polynucleotide contained in a mixture of polynucleotides such as a DNA or RNA library. As an alternative method for retrieving the partial sequences, screening hybridization is carried out under conditions 30 described above and using the formula for calculating  $T_m$ . Where fragments of 30 to 600 nucleotides are to be retrieved, the calculated  $T_m$  is corrected by subtracting (600/polynucleotide size in base pairs) and the stringency conditions are defined by a hybridization temperature that is 5 to 10°C below  $T_m$ . Where

oligonucleotides shorter than 20-30 bases are to be obtained, the formula for calculating the  $T_m$  is as follows:  $T_m = 4 \times (G+C) + 2 \times (A+T)$ . For example, an 18 nucleotide fragment of 50% G+C would have an approximate  $T_m$  of 54°C. Short peptides that are fragments of SEQ. ID Nos. 11 to 16 or their homologous sequences, are obtained directly by chemical synthesis (E. Gross and H. J. Meinhofer, 4 The Peptides: Analysis, Synthesis, Biology; Modern Techniques of Peptide Synthesis, John Wiley & Sons (1981), and M. Bodanzki, Principles of Peptide Synthesis, Springer-Verlag (1984)).

Useful polypeptide derivatives, e.g., polypeptide fragments, are designed using computer-assisted analysis of amino acid sequences. This identifies probable surface-exposed, antigenic regions (Hughes et al., 1992. Infect. Immun. 60(9):3497). An analysis of the 6 amino acid sequences contained in SEQ ID Nos: 11 to 16, based on the product of flexibility and hydrophobicity propensities using the program SEQSEE (Wishart DS, et al. "SEQSEE: a comprehensive program suite for protein sequence analysis." Comput Appl Biosci. 1994 Apr;10(2):121-32), reveal a number of potential B- and T-cell epitopes which may be used as a basis for selecting useful immunogenic fragments and variants. The results are shown in Figures 11 to 15. This analysis uses a reasonable combination of external surface features that is likely to be recognized by antibodies. Probable T-cell epitopes for HLA-A0201 MHC subclass were revealed by an algorithm written at Connaught Laboratories that emulates an approach developed at the NIH (Parker KC, et al. "Peptide binding to MHC class I molecules: implications for antigenic peptide prediction." Immunol Res 1995;14(1):34-57).

Epitopes which induce a protective T cell-dependent immune response are present throughout the length of the polypeptide. However, some epitopes may be masked by secondary and tertiary structures of the polypeptide. To reveal such masked epitopes large internal deletions are created which

remove much of the original protein structure and exposes the masked epitopes. Such internal deletions sometimes effects the additional advantage of removing immunodominant regions of high variability among strains. Polynucleotides encoding polypeptide 5 fragments and polypeptides having large internal deletions are constructed using standard methods (Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons Inc., 1994). Such methods include standard PCR, inverse PCR, restriction enzyme treatment of cloned DNA molecules, or the method of 10 Kunkel et al. (Kunkel et al. Proc. Natl. Acad. Sci. USA (1985) 82:448). Components for these methods and instructions for their use are readily available from various commercial sources such as Stratagene. Once the deletion mutants have been constructed, they are tested for their ability to prevent or 15 treat Chlamydia infection as described above.

As used herein, a fusion polypeptide is one that contains a polypeptide or a polypeptide derivative of the invention fused at the N- or C-terminal end to any other polypeptide (hereinafter referred to as a peptide tail). A simple way to 20 obtain such a fusion polypeptide is by translation of an in-frame fusion of the polynucleotide sequences, i.e., a hybrid gene. The hybrid gene encoding the fusion polypeptide is inserted into an expression vector which is used to transform or transfect a host cell. Alternatively, the polynucleotide 25 sequence encoding the polypeptide or polypeptide derivative is inserted into an expression vector in which the polynucleotide encoding the peptide tail is already present. Such vectors and instructions for their use are commercially available, e.g. the pMal-c2 or pMal-p2 system from New England Biolabs, in which the 30 peptide tail is a maltose binding protein, the glutathione-S-transferase system of Pharmacia, or the His-Tag system available from Novagen. These and other expression systems provide convenient means for further purification of polypeptides and derivatives of the invention.

An advantageous example of a fusion polypeptide is one where the polypeptide or homolog or fragment of the invention is fused to a polypeptide having adjuvant activity, such as subunit B of either cholera toxin or *E. coli* heat-labile toxin. Another 5 advantageous fusion is one where the polypeptide, homolog or fragment is fused to a strong T-cell epitope or B-cell epitope. Such an epitope may be one known in the art (e.g. the Hepatitis B virus core antigen, D.R. Millich et al., "Antibody production to the nucleocapsid and envelope of the Hepatitis B virus primed 10 by a single synthetic T cell site", *Nature*. 1987. 329:547-549), or one which has been identified in another polypeptide of the invention (Figures 11-15). Consistent with this aspect of the invention is a fusion polypeptide comprising T- or B-cell epitopes from one of SEQ ID Nos: 11 to 16 or its homolog or 15 fragment, wherein the epitopes are derived from multiple variants of said polypeptide or homolog or fragment, each variant differing from another in the location and sequence of its epitope within the polypeptide. Such a fusion is effective in the prevention and treatment of Chlamydia infection since it 20 optimizes the T- and B-cell response to the overall polypeptide, homolog or fragment.

To effect fusion, the polypeptide of the invention is fused to the N-, or preferably, to the C-terminal end of the polypeptide having adjuvant activity or T- or B-cell epitope. 25 Alternatively, a polypeptide fragment of the invention is inserted internally within the amino acid sequence of the polypeptide having adjuvant activity. The T- or B-cell epitope may also be inserted internally within the amino acid sequence of the polypeptide of the invention.

30 Consistent with the first aspect, the polynucleotides of the invention also encode hybrid precursor polypeptides containing heterologous signal peptides, which mature into polypeptides of the invention. By "heterologous signal peptide"

is meant a signal peptide that is not found in naturally-occurring precursors of polypeptides of the invention.

A polynucleotide molecule according to the invention, including RNA, DNA, or modifications or combinations thereof, have various applications. A DNA molecule is used, for example, (i) in a process for producing the encoded polypeptide in a recombinant host system, (ii) in the construction of vaccine vectors such as poxviruses, which are further used in methods and compositions for preventing and/or treating *Chlamydia* infection, (iii) as a vaccine agent (as well as an RNA molecule), in a naked form or formulated with a delivery vehicle and, (iv) in the construction of attenuated *Chlamydia* strains that can over-express a polynucleotide of the invention or express it in a non-toxic, mutated form.

Accordingly, a second aspect of the invention encompasses (i) an expression cassette containing a DNA molecule of the invention placed under the control of the elements required for expression, in particular under the control of an appropriate promoter; (ii) an expression vector containing an expression cassette of the invention; (iii) a procaryotic or eucaryotic cell transformed or transfected with an expression cassette and/or vector of the invention, as well as (iv) a process for producing a polypeptide or polypeptide derivative encoded by a polynucleotide of the invention, which involves culturing a procaryotic or eucaryotic cell transformed or transfected with an expression cassette and/or vector of the invention, under conditions that allow expression of the DNA molecule of the invention and, recovering the encoded polypeptide or polypeptide derivative from the cell culture.

A recombinant expression system is selected from procaryotic and eucaryotic hosts. Eucaryotic hosts include yeast cells (e.g., *Saccharomyces cerevisiae* or *Pichia pastoris*), mammalian cells (e.g., COS1, NIH3T3, or JEG3 cells), arthropods cells (e.g., *Spodoptera frugiperda* (SF9) cells), and plant

cells. A preferred expression system is a procaryotic host such as *E. coli*. Bacterial and eucaryotic cells are available from a number of different sources including commercial sources to those skilled in the art, e.g., the American Type Culture  
5 Collection (ATCC; Rockville, Maryland). Commercial sources of cells used for recombinant protein expression also provide instructions for usage of the cells.

The choice of the expression system depends on the features desired for the expressed polypeptide. For example, it  
10 may be useful to produce a polypeptide of the invention in a particular lipidated form or any other form.

One skilled in the art would readily understand that not all vectors and expression control sequences and hosts would be expected to express equally well the polynucleotides of this  
15 invention. With the guidelines described below, however, a selection of vectors, expression control sequences and hosts may be made without undue experimentation and without departing from the scope of this invention.

In selecting a vector, the host must be chosen that is  
20 compatible with the vector which is to exist and possibly replicate in it. Considerations are made with respect to the vector copy number, the ability to control the copy number, expression of other proteins such as antibiotic resistance. In selecting an expression control sequence, a number of variables  
25 are considered. Among the important variable are the relative strength of the sequence (e.g. the ability to drive expression under various conditions), the ability to control the sequence's function, compatibility between the polynucleotide to be expressed and the control sequence (e.g. secondary structures  
30 are considered to avoid hairpin structures which prevent efficient transcription). In selecting the host, unicellular hosts are selected which are compatible with the selected vector, tolerant of any possible toxic effects of the expressed product, able to secrete the expressed product efficiently if

such is desired, to be able to express the product in the desired conformation, to be easily scaled up, and to which ease of purification of the final product.

The choice of the expression cassette depends on the host system selected as well as the features desired for the expressed polypeptide. Typically, an expression cassette includes a promoter that is functional in the selected host system and can be constitutive or inducible; a ribosome binding site; a start codon (ATG) if necessary; a region encoding a signal peptide, e.g., a lipidation signal peptide; a DNA molecule of the invention; a stop codon; and optionally a 3' terminal region (translation and/or transcription terminator). The signal peptide encoding region is adjacent to the polynucleotide of the invention and placed in proper reading frame. The signal peptide-encoding region is homologous or heterologous to the DNA molecule encoding the mature polypeptide and is compatible with the secretion apparatus of the host used for expression. The open reading frame constituted by the DNA molecule of the invention, solely or together with the signal peptide, is placed under the control of the promoter so that transcription and translation occur in the host system.

Promoters and signal peptide encoding regions are widely known and available to those skilled in the art and include, for example, the promoter of *Salmonella typhimurium* (and derivatives) that is inducible by arabinose (promoter araB) and is functional in Gram-negative bacteria such as *E. coli* (as described in U.S. Patent No. 5,028,530 and in Cagnon et al., (Cagnon et al., Protein Engineering (1991) 4(7):843)); the promoter of the gene of bacteriophage T7 encoding RNA polymerase, that is functional in a number of *E. coli* strains expressing T7 polymerase (described in U.S. Patent No. 4,952,496); OspA lipidation signal peptide ; and RlpB lipidation signal peptide (Takase et al., J. Bact. (1987) 169:5692).



The expression cassette is typically part of an expression vector, which is selected for its ability to replicate in the chosen expression system. Expression vectors (e.g., plasmids or viral vectors) can be chosen, for example, 5 from those described in Pouwels et al. (Cloning Vectors: A Laboratory Manual 1985, Supp. 1987). Suitable expression vectors can be purchased from various commercial sources.

Methods for transforming/transfecting host cells with expression vectors are well-known in the art and depend on the 10 host system selected as described in Ausubel et al., (Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons Inc., 1994).

Upon expression, a recombinant polypeptide of the invention (or a polypeptide derivative) is produced and remains 15 in the intracellular compartment, is secreted/excreted in the extracellular medium or in the periplasmic space, or is embedded in the cellular membrane. The polypeptide is recovered in a substantially purified form from the cell extract or from the supernatant after centrifugation of the recombinant cell 20 culture. Typically, the recombinant polypeptide is purified by antibody-based affinity purification or by other well-known methods that can be readily adapted by a person skilled in the art, such as fusion of the polynucleotide encoding the polypeptide or its derivative to a small affinity binding 25 domain. Antibodies useful for purifying by immunoaffinity the polypeptides of the invention are obtained as described below.

A polynucleotide of the invention can also be useful as a vaccine. There are two major routes, either using a viral or bacterial host as gene delivery vehicle (live vaccine vector) or 30 administering the gene in a free form, e.g., inserted into a plasmid. Therapeutic or prophylactic efficacy of a polynucleotide of the invention is evaluated as described below.

Accordingly, a third aspect of the invention provides (i) a vaccine vector such as a poxvirus, containing a DNA molecule

of the invention, placed under the control of elements required for expression; (ii) a composition of matter comprising a vaccine vector of the invention, together with a diluent or carrier; specifically (iii) a pharmaceutical composition  
5 containing a therapeutically or prophylactically effective amount of a vaccine vector of the invention; (iv) a method for inducing an immune response against *Chlamydia* in a mammal (e.g., a human; alternatively, the method can be used in veterinary applications for treating or preventing *Chlamydia* infection of  
10 animals, e.g., cats or birds), which involves administering to the mammal an immunogenically effective amount of a vaccine vector of the invention to elicit a protective or therapeutic immune response to *Chlamydia* ; and particularly, (v) a method for preventing and/or treating a *Chlamydia* (e.g.,  
15 *C. trachomatis*, *C. psittaci*, *C. pneumonia*, *C. pecorum*) infection, which involves administering a prophylactic or therapeutic amount of a vaccine vector of the invention to an infected individual. Additionally, the third aspect of the invention encompasses the use of a vaccine vector of the  
20 invention in the preparation of a medicament for preventing and/or treating *Chlamydia* infection.

As used herein, a vaccine vector expresses one or several polypeptides or derivatives of the invention. The vaccine vector may express additionally a cytokine, such as interleukin-  
25 2 (IL-2) or interleukin-12 (IL-12), that enhances the immune response (adjuvant effect). It is understood that each of the components to be expressed is placed under the control of elements required for expression in a mammalian cell.

Consistent with the third aspect of the invention is a  
30 composition comprising several vaccine vectors, each of them capable of expressing a polypeptide or derivative of the invention. A composition may also comprise a vaccine vector capable of expressing an additional *Chlamydia* antigen, or a

subunit, fragment, homolog, mutant, or derivative thereof, optionally together with a cytokine such as IL-2 or IL-12.

Vaccination methods for treating or preventing infection in a mammal comprises use of a vaccine vector of the invention  
5 to be administered by any conventional route, particularly to a mucosal (e.g., ocular, intranasal, oral, gastric, pulmonary, intestinal, rectal, vaginal, or urinary tract) surface or via the parenteral (e.g., subcutaneous, intradermal, intramuscular, intravenous, or intraperitoneal) route. Preferred routes depend  
10 upon the choice of the vaccine vector. Treatment may be effected in a single dose or repeated at intervals. The appropriate dosage depends on various parameters understood by skilled artisans such as the vaccine vector itself, the route of administration or the condition of the mammal to be vaccinated  
15 (weight, age and the like).

Live vaccine vectors available in the art include viral vectors such as adenoviruses and poxviruses as well as bacterial vectors, e.g., *Shigella*, *Salmonella*, *Vibrio cholerae*, *Lactobacillus*, Bacille bilié de Calmette-Guérin (BCG), and  
20 *Streptococcus*.

An example of an adenovirus vector, as well as a method for constructing an adenovirus vector capable of expressing a DNA molecule of the invention, are described in U.S. Patent No. 4,920,209. Poxvirus vectors include vaccinia and canary pox  
25 virus, described in U.S. Patent No. 4,722,848 and U.S. Patent No. 5,364,773, respectively. (Also see, e.g., Tartaglia et al., Virology (1992) 188:217) for a description of a vaccinia virus vector and Taylor et al, Vaccine (1995) 13:539 for a reference of a canary pox.) Poxvirus vectors capable of expressing a  
30 polynucleotide of the invention are obtained by homologous recombination as described in Kieny et al., Nature (1984) 312:163 so that the polynucleotide of the invention is inserted in the viral genome under appropriate conditions for expression in mammalian cells. Generally, the dose of vaccine viral

vector, for therapeutic or prophylactic use, can be of from about  $1 \times 10^4$  to about  $1 \times 10^{11}$ , advantageously from about  $1 \times 10^7$  to about  $1 \times 10^{10}$ , preferably of from about  $1 \times 10^7$  to about  $1 \times 10^9$  plaque-forming units per kilogram. Preferably, viral vectors  
5 are administered parenterally; for example, in 3 doses, 4 weeks apart. It is preferable to avoid adding a chemical adjuvant to a composition containing a viral vector of the invention and thereby minimizing the immune response to the viral vector itself.

10 Non-toxicogenic *Vibrio cholerae* mutant strains that are useful as a live oral vaccine are known. Mekalanos et al., Nature (1983) 306:551 and U.S. Patent No. 4,882,278 describe strains which have a substantial amount of the coding sequence of each of the two *ctxA* alleles deleted so that no functional  
15 *cholerae* toxin is produced. WO 92/11354 describes a strain in which the *irgA* locus is inactivated by mutation; this mutation can be combined in a single strain with *ctxA* mutations. WO 94/01533 describes a deletion mutant lacking functional *ctxA* and *attRS1* DNA sequences. These mutant strains are genetically  
20 engineered to express heterologous antigens, as described in WO 94/19482. An effective vaccine dose of a *Vibrio cholerae* strain capable of expressing a polypeptide or polypeptide derivative encoded by a DNA molecule of the invention contains about  $1 \times 10^5$  to about  $1 \times 10^9$ , preferably about  $1 \times 10^6$  to about  $1 \times 10^8$ ,  
25 viable bacteria in a volume appropriate for the selected route of administration. Preferred routes of administration include all mucosal routes; most preferably, these vectors are administered intranasally or orally.

Attenuated *Salmonella typhimurium* strains, genetically  
30 engineered for recombinant expression of heterologous antigens or not, and their use as oral vaccines are described in Nakayama et al. (Bio/Technology (1988) 6:693) and WO 92/11361. Preferred routes of administration include all mucosal routes;

most preferably, these vectors are administered intranasally or orally.

Other bacterial strains used as vaccine vectors in the context of the present invention are described for *Shigella*  
5 *flexneri* in High et al., EMBO (1992) 11:1991 and Sizemore et al., Science (1995) 270:299; for *Streptococcus gordonii* in Medaglini et al., Proc. Natl. Acad. Sci. USA (1995) 92:6868; and for Bacille Calmette Guerin in Flynn J.L., Cell. Mol. Biol. (1994) 40 (suppl. I):31, WO 88/06626, WO 90/00594, WO 91/13157,  
10 WO 92/01796, and WO 92/21376.

In bacterial vectors, the polynucleotide of the invention is inserted into the bacterial genome or remains in a free state as part of a plasmid.

The composition comprising a vaccine bacterial vector of  
15 the present invention may further contain an adjuvant. A number of adjuvants are known to those skilled in the art. Preferred adjuvants as provided below.

Accordingly, a fourth aspect of the invention provides  
(i) a composition of matter comprising a polynucleotide of the  
20 invention, together with a diluent or carrier; (ii) a pharmaceutical composition comprising a therapeutically or prophylactically effective amount of a polynucleotide of the invention; (iii) a method for inducing an immune response against *Chlamydia* in a mammal by administration of an  
25 immunogenically effective amount of a polynucleotide of the invention to elicit a protective immune response to *Chlamydia*; and particularly, (iv) a method for preventing and/or treating a *Chlamydia* (e.g., *C. trachomatis*, *C. psittaci*, *C. pneumoniae*, or *C. pecorum*) infection, by administering a prophylactic or  
30 therapeutic amount of a polynucleotide of the invention to an infected individual. Additionally, the fourth aspect of the invention encompasses the use of a polynucleotide of the invention in the preparation of a medicament for preventing and/or treating *Chlamydia* infection. A preferred use includes

the use of a DNA molecule placed under conditions for expression in a mammalian cell, especially in a plasmid that is unable to replicate in mammalian cells and to substantially integrate in a mammalian genome.

5           Use of the polynucleotides of the invention include their administration to a mammal as a vaccine, for therapeutic or prophylactic purposes. Such polynucleotides are used in the form of DNA as part of a plasmid that is unable to replicate in a mammalian cell and unable to integrate into the mammalian  
10 genome. Typically, such a DNA molecule is placed under the control of a promoter suitable for expression in a mammalian cell. The promoter functions either ubiquitously or tissue-specifically. Examples of non-tissue specific promoters include the early Cytomegalovirus (CMV) promoter (described in U.S.  
15 Patent No. 4,168,062) and the Rous Sarcoma Virus promoter (described in Norton & Coffin, Molec. Cell Biol. (1985) 5:281). An example of a tissue-specific promoter is the desmin promoter which drives expression in muscle cells (Li et al., Gene (1989) 78:243, Li & Paulin, J. Biol. Chem. (1991) 266:6562 and Li &  
20 Paulin, J. Biol. Chem. (1993) 268:10403). Use of promoters is well-known to those skilled in the art. Useful vectors are described in numerous publications, specifically WO 94/21797 and Hartikka et al., Human Gene Therapy (1996) 7:1205.

Polynucleotides of the invention which are used as a  
25 vaccine encode either a precursor or a mature form of the corresponding polypeptide. In the precursor form, the signal peptide is either homologous or heterologous. In the latter case, a eucaryotic leader sequence such as the leader sequence of the tissue-type plasminogen factor (tPA) is preferred.

30           As used herein, a composition of the invention contains one or several polynucleotides with optionally at least one additional polynucleotide encoding another *Chlamydia* antigen such as urease subunit A, B, or both, or a fragment, derivative, mutant, or analog thereof. The composition may also contain an

additional polynucleotide encoding a cytokine, such as interleukin-2 (IL-2) or interleukin-12 (IL-12) so that the immune response is enhanced. These additional polynucleotides are placed under appropriate control for expression.

5 Advantageously, DNA molecules of the invention and/or additional DNA molecules to be included in the same composition, are present in the same plasmid.

Standard techniques of molecular biology for preparing and purifying polynucleotides are used in the preparation of  
10 polynucleotide therapeutics of the invention. For use as a vaccine, a polynucleotide of the invention is formulated according to various methods outlined below.

One method utilizes the polynucleotide in a naked form, free of any delivery vehicles. Such a polynucleotide is simply  
15 diluted in a physiologically acceptable solution such as sterile saline or sterile buffered saline, with or without a carrier. When present, the carrier preferably is isotonic, hypotonic, or weakly hypertonic, and has a relatively low ionic strength, such as provided by a sucrose solution, e.g., a solution containing  
20 20% sucrose.

An alternative method utilizes the polynucleotide in association with agents that assist in cellular uptake. Examples of such agents are (i) chemicals that modify cellular permeability, such as bupivacaine (see, e.g., WO 94/16737), (ii)  
25 liposomes for encapsulation of the polynucleotide, or (iii) cationic lipids or silica, gold, or tungsten microparticles which associate themselves with the polynucleotides.

Anionic and neutral liposomes are well-known in the art  
30 (see, e.g., Liposomes: A Practical Approach, RPC New Ed, IRL press (1990), for a detailed description of methods for making liposomes) and are useful for delivering a large range of products, including polynucleotides. Cationic lipids are also known in the art and are commonly used for gene delivery. Such

lipids include Lipofectin<sup>TM</sup> also known as DOTMA (N-[1-(2,3-dioleyloxy)propyl]-N,N,N-trimethylammonium chloride), DOTAP (1,2-bis(oleyloxy)-3-(trimethylammonio)propane), DDAB (dimethyldioctadecylammonium bromide), DOGS (dioctadecylamidologlycyl spermine) and cholesterol derivatives such as DC-Chol (3 beta-(N-(N',N'-dimethyl aminomethane)-carbamoyl) cholesterol). A description of these cationic lipids can be found in EP 187,702, WO 90/11092, U.S. Patent No. 5,283,185, WO 91/15501, WO 95/26356, and U.S. Patent No. 5,527,928. Cationic lipids for gene delivery are preferably used in association with a neutral lipid such as DOPE (dioleoyl phosphatidylethanolamine), as described in WO 90/11092 as an example.

Formulation containing cationic liposomes may optionally contain other transfection-facilitating compounds. A number of them are described in WO 93/18759, WO 93/19768, WO 94/25608, and WO 95/02397. They include spermine derivatives useful for facilitating the transport of DNA through the nuclear membrane (see, for example, WO 93/18759) and membrane-permeabilizing compounds such as GALA, Gramicidine S, and cationic bile salts (see, for example, WO 93/19768).

Gold or tungsten microparticles are used for gene delivery, as described in WO 91/00359, WO 93/17706, and Tang et al. Nature (1992) 356:152. The microparticle-coated polynucleotide is injected via intradermal or intraepidermal routes using a needleless injection device ("gene gun"), such as those described in U.S. Patent No. 4,945,050, U.S. Patent No. 5,015,580, and WO 94/24263.

The amount of DNA to be used in a vaccine recipient depends, e.g., on the strength of the promoter used in the DNA construct, the immunogenicity of the expressed gene product, the condition of the mammal intended for administration (e.g., the weight, age, and general health of the mammal), the mode of administration, and the type of formulation. In general, a



therapeutically or prophylactically effective dose from about 1 µg to about 1 mg, preferably, from about 10 µg to about 800 µg and, more preferably, from about 25 µg to about 250 µg, can be administered to human adults. The administration can be  
5 achieved in a single dose or repeated at intervals.

The route of administration is any conventional route used in the vaccine field. As general guidance, a polynucleotide of the invention is administered via a mucosal surface, e.g., an ocular, intranasal, pulmonary, oral,  
10 intestinal, rectal, vaginal, and urinary tract surface; or via a parenteral route, e.g., by an intravenous, subcutaneous, intraperitoneal, intradermal, intraepidermal, or intramuscular route. The choice of administration route depends on the formulation that is selected. A polynucleotide formulated in  
15 association with bupivacaine is advantageously administered into muscles. When a neutral or anionic liposome or a cationic lipid, such as DOTMA or DC-Chol, is used, the formulation can be advantageously injected via intravenous, intranasal (aerosolization), intramuscular, intradermal, and subcutaneous  
20 routes. A polynucleotide in a naked form can advantageously be administered via the intramuscular, intradermal, or subcutaneous routes.

Although not absolutely required, such a composition can also contain an adjuvant. If so, a systemic adjuvant that does  
25 not require concomitant administration in order to exhibit an adjuvant effect is preferable such as, e.g., QS21, which is described in U.S. Patent No. 5,057,546.

The sequence information provided in the present application enables the design of specific nucleotide probes and  
30 primers that are used for diagnostic purposes. Accordingly, a fifth aspect of the invention provides a nucleotide probe or primer having a sequence found in or derived by degeneracy of the genetic code from a sequence shown in any one of SEQ ID Nos: 1 to 10.

The term "probe" as used in the present application refers to DNA (preferably single stranded) or RNA molecules (or modifications or combinations thereof) that hybridize under the stringent conditions, as defined above, to nucleic acid molecules having SEQ ID Nos: 1 to 10 or to sequences homologous to SEQ ID Nos: 1 to 10, or to their complementary or anti-sense sequences. Generally, probes are significantly shorter than full-length sequences. Such probes contain from about 5 to about 100, preferably from about 10 to about 80, nucleotides. In particular, probes have sequences that are at least 75%, preferably at least 85%, more preferably 95% homologous to a portion of any of SEQ ID Nos: 1 to 10 or that are complementary to such sequences. Probes may contain modified bases such as inosine, methyl-5-deoxycytidine, deoxyuridine, dimethylamino-5-deoxyuridine, or diamino-2, 6-purine. Sugar or phosphate residues may also be modified or substituted. For example, a deoxyribose residue may be replaced by a polyamide (Nielsen et al., Science (1991) 254:1497) and phosphate residues may be replaced by ester groups such as diphosphate, alkyl, arylphosphonate and phosphorothioate esters. In addition, the 2'-hydroxyl group on ribonucleotides may be modified by including such groups as alkyl groups.

Probes of the invention are used in diagnostic tests, as capture or detection probes. Such capture probes are conventionally immobilized on a solid support, directly or indirectly, by covalent means or by passive adsorption. A detection probe is labelled by a detection marker selected from: radioactive isotopes, enzymes such as peroxidase, alkaline phosphatase, and enzymes able to hydrolyze a chromogenic, fluorogenic, or luminescent substrate, compounds that are chromogenic, fluorogenic, or luminescent, nucleotide base analogs, and biotin.

Probes of the invention are used in any conventional hybridization technique, such as dot blot (Maniatis et al.,

Molecular Cloning: A Laboratory Manual (1982) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York), Southern blot (Southern, J. Mol. Biol. (1975) 98:503), northern blot (identical to Southern blot with the exception that RNA is used as a target), or the sandwich technique (Dunn et al., Cell (1977) 12:23). The latter technique involves the use of a specific capture probe and/or a specific detection probe with nucleotide sequences that at least partially differ from each other.

10       A primer is a probe of usually about 10 to about 40 nucleotides that is used to initiate enzymatic polymerization of DNA in an amplification process (e.g., PCR), in an elongation process, or in a reverse transcription method. Primers used in diagnostic methods involving PCR are labeled by methods known in  
15 the art.

As described herein, the invention also encompasses (i) a reagent comprising a probe of the invention for detecting and/or identifying the presence of *Chlamydia* in a biological material; (ii) a method for detecting and/or identifying the presence of  
20 *Chlamydia* in a biological material, in which (a) a sample is recovered or derived from the biological material, (b) DNA or RNA is extracted from the material and denatured, and (c) exposed to a probe of the invention, for example, a capture, detection probe or both, under stringent hybridization  
25 conditions, such that hybridization is detected; and (iii) a method for detecting and/or identifying the presence of *Chlamydia* in a biological material, in which (a) a sample is recovered or derived from the biological material, (b) DNA is extracted therefrom, (c) the extracted DNA is primed with at  
30 least one, and preferably two, primers of the invention and amplified by polymerase chain reaction, and (d) the amplified DNA fragment is produced.

It is apparent that disclosure of polynucleotide sequences of SEQ ID Nos: 1 to 10, their homolog, and partial

sequences of either enable their corresponding amino acid sequences. Accordingly, a sixth aspect of the invention features a substantially purified polypeptide or polypeptide derivative having an amino acid sequence encoded by a  
5 polynucleotide of the invention.

A "substantially purified polypeptide" as used herein is defined as a polypeptide that is separated from the environment in which it naturally occurs and/or that is free of the majority of the polypeptides that are present in the environment in which  
10 it was synthesized. For example, a substantially purified polypeptide is free from cytoplasmic polypeptides. Those skilled in the art would readily understand that the polypeptides of the invention may be purified from a natural source, i.e., a *Chlamydia* strain, or produced by recombinant  
15 means.

Consistent with the sixth aspect of the invention are polypeptides, homologs or fragments which are modified or treated to enhance their immunogenicity in the target animal, in whom the polypeptide, homolog or fragments are intended to  
20 confer protection against *Chlamydia*. Such modifications or treatments include: amino acid substitutions with an amino acid derivative such as 3-methylhistidine, 4-hydroxyproline, 5-hydroxylysine etc., modifications or deletions which are carried out after preparation of the polypeptide, homolog or fragment,  
25 such as the modification of free amino, carboxyl or hydroxyl side groups of the amino acids.

Identification of homologous polypeptides or polypeptide derivatives encoded by polynucleotides of the invention which have specific antigenicity is achieved by screening for cross-  
30 reactivity with an antiserum raised against the polypeptide of reference having an amino acid sequence of any one of SEQ ID Nos: 11 to 16. The procedure is as follows: a monospecific hyperimmune antiserum is raised against a purified reference polypeptide, a fusion polypeptide (for example, an expression

product of MBP, GST, or His-tag systems), or a synthetic peptide predicted to be antigenic. Where an antiserum is raised against a fusion polypeptide, two different fusion systems are employed. Specific antigenicity can be determined according to a number of 5 methods, including Western blot (Towbin et al., Proc. Natl. Acad. Sci. USA (1979) 76:4350), dot blot, and ELISA, as described below.

In a Western blot assay, the product to be screened, either as a purified preparation or a total *E. coli* extract, is 10 submitted to SDS-Page electrophoresis as described by Laemmli (Nature (1970) 227:680). After transfer to a nitrocellulose membrane, the material is further incubated with the monospecific hyperimmune antiserum diluted in the range of dilutions from about 1:5 to about 1:5000, preferably from about 15 1:100 to about 1:500. Specific antigenicity is shown once a band corresponding to the product exhibits reactivity at any of the dilutions in the above range.

In an ELISA assay, the product to be screened is preferably used as the coating antigen. A purified preparation 20 is preferred, although a whole cell extract can also be used. Briefly, about 100  $\mu$ l of a preparation at about 10  $\mu$ g protein/ml are distributed into wells of a 96-well polycarbonate ELISA plate. The plate is incubated for 2 hours at 37°C then overnight at 4°C. The plate is washed with phosphate buffer 25 saline (PBS) containing 0.05% Tween 20 (PBS/Tween buffer). The wells are saturated with 250  $\mu$ l PBS containing 1% bovine serum albumin (BSA) to prevent non-specific antibody binding. After 1 hour incubation at 37°C, the plate is washed with PBS/Tween buffer. The antiserum is serially diluted in PBS/Tween buffer 30 containing 0.5% BSA. 100  $\mu$ l of dilutions are added per well. The plate is incubated for 90 minutes at 37°C, washed and evaluated according to standard procedures. For example, a goat anti-rabbit peroxidase conjugate is added to the wells when specific antibodies were raised in rabbits. Incubation is

carried out for 90 minutes at 37°C and the plate is washed. The reaction is developed with the appropriate substrate and the reaction is measured by colorimetry (absorbance measured spectrophotometrically). Under the above experimental conditions, a positive reaction is shown by O.D. values greater than a non immune control serum.

In a dot blot assay, a purified product is preferred, although a whole cell extract can also be used. Briefly, a solution of the product at about 100 µg/ml is serially two-fold diluted in 50 mM Tris-HCl (pH 7.5). 100 µl of each dilution are applied to a nitrocellulose membrane 0.45 µm set in a 96-well dot blot apparatus (Biorad). The buffer is removed by applying vacuum to the system. Wells are washed by addition of 50 mM Tris-HCl (pH 7.5) and the membrane is air-dried. The membrane is saturated in blocking buffer (50 mM Tris-HCl (pH 7.5) 0.15 M NaCl, 10 g/L skim milk) and incubated with an antiserum dilution from about 1:50 to about 1:5000, preferably about 1:500. The reaction is revealed according to standard procedures. For example, a goat anti-rabbit peroxidase conjugate is added to the wells when rabbit antibodies are used. Incubation is carried out 90 minutes at 37°C and the blot is washed. The reaction is developed with the appropriate substrate and stopped. The reaction is measured visually by the appearance of a colored spot, e.g., by colorimetry. Under the above experimental conditions, a positive reaction is shown once a colored spot is associated with a dilution of at least about 1:5, preferably of at least about 1:500.

Therapeutic or prophylactic efficacy of a polypeptide or derivative of the invention can be evaluated as described below. A seventh aspect of the invention provides (i) a composition of matter comprising a polypeptide of the invention together with a diluent or carrier; specifically (ii) a pharmaceutical composition containing a therapeutically or prophylactically effective amount of a polypeptide of the invention; (iii) a

method for inducing an immune response against *Chlamydia* in a mammal, by administering to the mammal an immunogenically effective amount of a polypeptide of the invention to elicit a protective immune response to *Chlamydia*; and particularly, (iv) 5 a method for preventing and/or treating a *Chlamydia* (e.g., *C. trachomatis*, *C. psittaci*, *C. pneumoniae*, or *C. pecorum*) infection, by administering a prophylactic or therapeutic amount of a polypeptide of the invention to an infected individual. Additionally, the seventh aspect of the invention encompasses 10 the use of a polypeptide of the invention in the preparation of a medicament for preventing and/or treating *Chlamydia* infection.

As used herein, the immunogenic compositions of the invention are administered by conventional routes known the vaccine field, in particular to a mucosal (e.g., ocular, 15 intranasal, pulmonary, oral, gastric, intestinal, rectal, vaginal, or urinary tract) surface or via the parenteral (e.g., subcutaneous, intradermal, intramuscular, intravenous, or intraperitoneal) route. The choice of administration route depends upon a number of parameters, such as the adjuvant 20 associated with the polypeptide. If a mucosal adjuvant is used, the intranasal or oral route is preferred. If a lipid formulation or an aluminum compound is used, the parenteral route is preferred with the sub-cutaneous or intramuscular route being most preferred. The choice also depends upon the nature 25 of the vaccine agent. For example, a polypeptide of the invention fused to CTB or LTB is best administered to a mucosal surface.

As used herein, the composition of the invention contains one or several polypeptides or derivatives of the invention. 30 The composition optionally contains at least one additional *Chlamydia* antigen, or a subunit, fragment, homolog, mutant, or derivative thereof.

For use in a composition of the invention, a polypeptide or derivative thereof is formulated into or with liposomes,

preferably neutral or anionic liposomes, microspheres, ISCOMS, or virus-like-particles (VLPs) to facilitate delivery and/or enhance the immune response. These compounds are readily available to one skilled in the art; for example, see Liposomes: 5 A Practical Approach, RPC New Ed, IRL press (1990).

Adjuvants other than liposomes and the like are also used and are known in the art. Adjuvants may protect the antigen from rapid dispersal by sequestering it in a local deposit, or they may contain substances that stimulate the host to secrete 10 factors that are chemotactic for macrophages and other components of the immune system. An appropriate selection can conventionally be made by those skilled in the art, for example, from those described below (see the eleventh aspect of the invention).

15 Treatment is achieved in a single dose or repeated as necessary at intervals, as can be determined readily by one skilled in the art. For example, a priming dose is followed by three booster doses at weekly or monthly intervals. An appropriate dose depends on various parameters including the 20 recipient (e.g., adult or infant), the particular vaccine antigen, the route and frequency of administration, the presence/absence or type of adjuvant, and the desired effect (e.g., protection and/or treatment), as can be determined by one skilled in the art. In general, a vaccine antigen of the 25 invention is administered by a mucosal route in an amount from about 10 µg to about 500 mg, preferably from about 1 mg to about 200 mg. For the parenteral route of administration, the dose usually does not exceed about 1 mg, preferably about 100 µg.

When used as vaccine agents, polynucleotides and 30 polypeptides of the invention may be used sequentially as part of a multistep immunization process. For example, a mammal is initially primed with a vaccine vector of the invention such as a pox virus, e.g., via the parenteral route, and then boosted twice with the polypeptide encoded by the vaccine vector, e.g.,



via the mucosal route. In another example, liposomes associated with a polypeptide or derivative of the invention is also used for priming, with boosting being carried out mucosally using a soluble polypeptide or derivative of the invention in combination with a mucosal adjuvant (e.g., LT).

A polypeptide derivative of the invention is also used in accordance with the seventh aspect as a diagnostic reagent for detecting the presence of anti-*Chlamydia* antibodies, e.g., in a blood sample. Such polypeptides are about 5 to about 80, preferably about 10 to about 50 amino acids in length. They are either labeled or unlabeled, depending upon the diagnostic method. Diagnostic methods involving such a reagent are described below.

Upon expression of a DNA molecule of the invention, a polypeptide or polypeptide derivative is produced and purified using known laboratory techniques. As described above, the polypeptide or polypeptide derivative may be produced as a fusion protein containing a fused tail that facilitates purification. The fusion product is used to immunize a small mammal, e.g., a mouse or a rabbit, in order to raise antibodies against the polypeptide or polypeptide derivative (monospecific antibodies). Accordingly, an eighth aspect of the invention provides a monospecific antibody that binds to a polypeptide or polypeptide derivative of the invention.

By "monospecific antibody" is meant an antibody that is capable of reacting with a unique naturally-occurring *Chlamydia* polypeptide. An antibody of the invention is either polyclonal or monoclonal. Monospecific antibodies may be recombinant, e.g., chimeric (e.g., constituted by a variable region of murine origin associated with a human constant region), humanized (a human immunoglobulin constant backbone together with hypervariable region of animal, e.g., murine, origin), and/or single chain. Both polyclonal and monospecific antibodies may also be in the form of immunoglobulin fragments, e.g., F(ab)'2

or Fab fragments. The antibodies of the invention are of any isotype, e.g., IgG or IgA, and polyclonal antibodies are of a single isotype or a mixture of isotypes.

Antibodies against the polypeptides, homologs or  
5 fragments of the present invention are generated by immunization of a mammal with a composition comprising said polypeptide, homolog or fragment. Such antibodies may be polyclonal or monoclonal. Methods to produce polyclonal or monoclonal antibodies are well known in the art. For a review, see  
10 "Antibodies, A Laboratory Manual, Cold Spring Harbor Laboratory, Eds. E. Harlow and D. Lane (1988), and D.E. Yelton et al., 1981. Ann. Rev. Biochem. 50:657-680. For monoclonal antibodies, see Kohler and Milstein (1975) Nature. 256:495-497.

The antibodies of the invention, which are raised to a  
15 polypeptide or polypeptide derivative of the invention, are produced and identified using standard immunological assays, e.g., Western blot analysis, dot blot assay, or ELISA (see, e.g., Coligan et al., Current Protocols in Immunology (1994) John Wiley & Sons, Inc., New York, NY). The antibodies are used  
20 in diagnostic methods to detect the presence of a *Chlamydia* antigen in a sample, such as a biological sample. The antibodies are also used in affinity chromatography for purifying a polypeptide or polypeptide derivative of the invention. As is discussed further below, such antibodies may  
25 be used in prophylactic and therapeutic passive immunization methods.

Accordingly, a ninth aspect of the invention provides  
(i) a reagent for detecting the presence of *Chlamydia* in a biological sample that contains an antibody, polypeptide, or  
30 polypeptide derivative of the invention; and (ii) a diagnostic method for detecting the presence of *Chlamydia* in a biological sample, by contacting the biological sample with an antibody, a polypeptide, or a polypeptide derivative of the invention, such that an immune complex is formed, and by detecting such complex

to indicate the presence of *Chlamydia* in the sample or the organism from which the sample is derived.

Those skilled in the art will readily understand that the immune complex is formed between a component of the sample and the antibody, polypeptide, or polypeptide derivative, whichever is used, and that any unbound material is removed prior to detecting the complex. It is understood that a polypeptide reagent is useful for detecting the presence of anti-*Chlamydia* antibodies in a sample, e.g., a blood sample, while an antibody of the invention is used for screening a sample, such as a gastric extract or biopsy, for the presence of *Chlamydia* polypeptides.

For diagnostic applications, the reagent (i.e., the antibody, polypeptide, or polypeptide derivative of the invention) is either in a free state or immobilized on a solid support, such as a tube, a bead, or any other conventional support used in the field. Immobilization is achieved using direct or indirect means. Direct means include passive adsorption (non-covalent binding) or covalent binding between the support and the reagent. By "indirect means" is meant that an anti-reagent compound that interacts with a reagent is first attached to the solid support. For example, if a polypeptide reagent is used, an antibody that binds to it can serve as an anti-reagent, provided that it binds to an epitope that is not involved in the recognition of antibodies in biological samples. Indirect means may also employ a ligand-receptor system, for example, where a molecule such as a vitamin is grafted onto the polypeptide reagent and the corresponding receptor immobilized on the solid phase. This is illustrated by the biotin-streptavidin system. Alternatively, a peptide tail is added chemically or by genetic engineering to the reagent and the grafted or fused product immobilized by passive adsorption or covalent linkage of the peptide tail.

Such diagnostic agents may be included in a kit which also comprises instructions for use. The reagent are labeled with a detection means which allows for the detection of the reagent when it is bound to its target. The detection means may  
5 be a fluorescent agent such as fluorescein isocyanate or fluorescein isothiocyanate, or an enzyme such as horse radish peroxidase or luciferase or alkaline phosphatase, or a radioactive element such as  $^{125}\text{I}$  or  $^{51}\text{Cr}$ .

Accordingly, a tenth aspect of the invention provides a  
10 process for purifying, from a biological sample, a polypeptide or polypeptide derivative of the invention, which involves carrying out antibody-based affinity chromatography with the biological sample, wherein the antibody is a monospecific antibody of the invention.

15 For use in a purification process of the invention, the antibody is either polyclonal or monospecific, and preferably is of the IgG type. Purified IgGs is prepared from an antiserum using standard methods (see, e.g., Coligan et al., Current Protocols in Immunology (1994) John Wiley & Sons, Inc., New  
20 York, NY). Conventional chromatography supports, as well as standard methods for grafting antibodies, are described in, e.g., Antibodies: A Laboratory Manual, D. Lane, E. Harlow, Eds. (1988) and outlined below.

Briefly, a biological sample, such as an *C. pneumoniae*  
25 extract preferably in a buffer solution, is applied to a chromatography material, preferably equilibrated with the buffer used to dilute the biological sample so that the polypeptide or polypeptide derivative of the invention (i.e., the antigen) is allowed to adsorb onto the material. The chromatography  
30 material, such as a gel or a resin coupled to an antibody of the invention, is in either a batch form or a column. The unbound components are washed off and the antigen is then eluted with an appropriate elution buffer, such as a glycine buffer or a buffer containing a chaotropic agent, e.g., guanidine HCl, or high salt

concentration (e.g., 3 M  $\text{MgCl}_2$ ). Eluted fractions are recovered and the presence of the antigen is detected, e.g., by measuring the absorbance at 280 nm.

An eleventh aspect of the invention provides (i) a  
5 composition of matter comprising a monospecific antibody of the invention, together with a diluent or carrier; (ii) a pharmaceutical composition comprising a therapeutically or prophylactically effective amount of a monospecific antibody of the invention, and (iii) a method for treating or preventing a  
10 *Chlamydia* (e.g., *C. trachomatis*, *C. psittaci*, *C. pneumoniae* or *C. pecorum*) infection, by administering a therapeutic or prophylactic amount of a monospecific antibody of the invention to an infected individual. Additionally, the eleventh aspect of the invention encompasses the use of a monospecific antibody of  
15 the invention in the preparation of a medicament for treating or preventing *Chlamydia* infection.

The monospecific antibody is either polyclonal or monoclonal, preferably of the IgA isotype (predominantly). In passive immunization, the antibody is administered to a mucosal  
20 surface of a mammal, e.g., the gastric mucosa, e.g., orally or intragastrically, advantageously, in the presence of a bicarbonate buffer. Alternatively, systemic administration, not requiring a bicarbonate buffer, is carried out. A monospecific antibody of the invention is administered as a single active  
25 component or as a mixture with at least one monospecific antibody specific for a different *Chlamydia* polypeptide. The amount of antibody and the particular regimen used are readily determined by one skilled in the art. For example, daily administration of about 100 to 1,000 mg of antibodies over one  
30 week, or three doses per day of about 100 to 1,000 mg of antibodies over two or three days, are effective regimens for most purposes.

Therapeutic or prophylactic efficacy are evaluated using standard methods in the art, e.g., by measuring induction of a

mucosal immune response or induction of protective and/or therapeutic immunity, using, e.g., the *C. pneumoniae* mouse model. Those skilled in the art will readily recognize that the *C. pneumoniae* strain of the model may be replaced with another 5 *Chlamydia* strain. For example, the efficacy of DNA molecules and polypeptides from *C. pneumoniae* is preferably evaluated in a mouse model using *C. pneumoniae* strain. Protection is determined by comparing the degree of *Chlamydia* infection to that of a control group. Protection is shown when infection is 10 reduced by comparison to the control group. Such an evaluation is made for polynucleotides, vaccine vectors, polypeptides and derivatives thereof, as well as antibodies of the invention.

Adjuvants useful in any of the vaccine compositions described above are as follows.

15 Adjuvants for parenteral administration include aluminum compounds, such as aluminum hydroxide, aluminum phosphate, and aluminum hydroxy phosphate. The antigen is precipitated with, or adsorbed onto, the aluminum compound according to standard protocols. Other adjuvants, such as RIBI (ImmunoChem, Hamilton, 20 MT), is used in parenteral administration.

Adjuvants for mucosal administration include bacterial toxins, e.g., the cholera toxin (CT), the *E. coli* heat-labile toxin (LT), the *Clostridium difficile* toxin A and the pertussis toxin (PT), or combinations, subunits, toxoids, or mutants 25 thereof such as a purified preparation of native cholera toxin subunit B (CTB). Fragments, homologs, derivatives, and fusions to any of these toxins are also suitable, provided that they retain adjuvant activity. Preferably, a mutant having reduced toxicity is used. Suitable mutants are described, e.g., in WO 30 95/17211 (Arg-7-Lys CT mutant), WO 96/06627 (Arg-192-Gly LT mutant), and WO 95/34323 (Arg-9-Lys and Glu-129-Gly PT mutant). Additional LT mutants that are used in the methods and compositions of the invention include, e.g., Ser-63-Lys, Ala-69-Gly, Glu-110-Asp, and Glu-112-Asp mutants. Other adjuvants,

such as a bacterial monophosphoryl lipid A (MPLA) of, e.g., *E. coli*, *Salmonella minnesota*, *Salmonella typhimurium*, or *Shigella flexneri*; saponins, or polylactide glycolide (PLGA) microspheres, is also be used in mucosal administration.

5           Adjuvants useful for both mucosal and parenteral administrations include polyphosphazene (WO 95/02415), DC-chol (3 b-(N-(N',N'-dimethyl aminomethane)-carbamoyl) cholesterol; U.S. Patent No. 5,283,185 and WO 96/14831) and QS-21 (WO 88/09336).

10           Any pharmaceutical composition of the invention containing a polynucleotide, a polypeptide, a polypeptide derivative, or an antibody of the invention, is manufactured in a conventional manner. In particular, it is formulated with a pharmaceutically acceptable diluent or carrier, e.g., water or a  
15 saline solution such as phosphate buffer saline. In general, a diluent or carrier is selected on the basis of the mode and route of administration, and standard pharmaceutical practice. Suitable pharmaceutical carriers or diluents, as well as pharmaceutical necessities for their use in pharmaceutical  
20 formulations, are described in *Remington's Pharmaceutical Sciences*, a standard reference text in this field and in the USP/NF.

          The invention also includes methods in which *Chlamydia* infection are treated by oral administration of a *Chlamydia*  
25 polypeptide of the invention and a mucosal adjuvant, in combination with an antibiotic, an antacid, sucralfate, or a combination thereof. Examples of such compounds that can be administered with the vaccine antigen and the adjuvant are antibiotics, including, e.g., macrolides, tetracyclines, and  
30 derivatives thereof (specific examples of antibiotics that can be used include azithromycin or doxycyclin or immunomodulators such as cytokines or steroids). In addition, compounds containing more than one of the above-listed components coupled together, are used. The invention also includes compositions

for carrying out these methods, i.e., compositions containing a *Chlamydia* antigen (or antigens) of the invention, an adjuvant, and one or more of the above-listed compounds, in a pharmaceutically acceptable carrier or diluent.

5           Amounts of the above-listed compounds used in the methods and compositions of the invention are readily determined by one skilled in the art. Treatment/immunization schedules are also known and readily designed by one skilled in the art. For example, the non-vaccine components can be administered on days  
10 1-14, and the vaccine antigen + adjuvant can be administered on days 7, 14, 21, and 28.



CLAIMS

- 5 1. A nucleic acid molecule comprising a nucleic acid sequence which encodes a polypeptide selected from any of:
- (a) SEQ ID Nos: 11 to 16;
  - (b) an immunogenic fragment comprising at least 12 consecutive amino acids from a polypeptide of (a); and
  - 10 (c) a polypeptide of (a) or (b) which has been modified to improve its immunogenicity, wherein said modified polypeptide is at least 75% identical in amino acid sequence to the corresponding polypeptide of (a) or (b).
- 15
2. A nucleic acid molecule comprising a nucleic acid sequence selected from any of:
- (a) SEQ ID Nos: 1 to 10;
  - (b) a sequence which encodes a polypeptide encoded by any
  - 20 one of SEQ ID Nos: 1 to 10;
  - (c) a sequence comprising at least 38 consecutive nucleotides from any one of the nucleic acid sequences of (a) and (b); and
  - (d) a sequence which encodes a polypeptide which is at
  - 25 least 75% identical in amino acid sequence to any one of the polypeptides encoded by SEQ ID Nos: 1 to 10.

3. A nucleic acid molecule comprising a nucleic acid sequence which is anti-sense to the nucleic acid molecule of claim 1 or 2.
- 5 4. A nucleic acid molecule comprising a nucleic acid sequence which encodes a fusion protein, said fusion protein comprising a polypeptide encoded by a nucleic acid molecule according to claim 1 and an additional polypeptide.
- 10 5. The nucleic acid molecule of claim 4 wherein the additional polypeptide is a heterologous signal peptide.
6. The nucleic acid molecule of claim 4 wherein the additional polypeptide has adjuvant activity.
- 15 7. A nucleic acid molecule according to any one of claims 1 to 6, operatively linked to one or more expression control sequences.
- 20 8. A vaccine comprising at least one first nucleic acid according to any one of claims 1, 2, and 4 to 7 and a vaccine vector wherein each first nucleic acid is expressed as a polypeptide, the vaccine optionally comprising a second nucleic acid encoding an additional polypeptide which enhances the
- 25 immune response to the polypeptide expressed by said first nucleic acid.

9. The vaccine of claim 8 wherein the second nucleic acid encodes an additional *Chlamydia* polypeptide.

5 10. A pharmaceutical composition comprising a nucleic acid according to any one of claims 1 to 7 and a pharmaceutically acceptable carrier.

11. A pharmaceutical composition comprising a vaccine  
10 according to claim 8 or 9 and a pharmaceutically acceptable carrier.

12. A unicellular host transformed with the nucleic acid molecule of claim 7.

15

13. A nucleic acid probe of 5 to 100 nucleotides which hybridizes under stringent conditions to any one of nucleic acid molecules of SEQ ID Nos: 1 to 10, or to a homolog or complementary or anti-sense sequence of said nucleic acid  
20 molecule.

14. A primer of 10 to 40 nucleotides which hybridizes under stringent conditions to any one of nucleic acid molecules of SEQ ID Nos: 1 to 10, or to a homolog or complementary or  
25 anti-sense sequence of said nucleic acid molecule.

15. A polypeptide encoded by a nucleic acid sequence according to any one of claims 1, 2 and 4 to 7.

16. A polypeptide comprising an amino acid sequence  
5 selected from any of:

- (a) SEQ ID Nos: 11 to 16;
- (b) an immunogenic fragment comprising at least 12 consecutive amino acids from a polypeptide of (a); and
- (c) a polypeptide of (a) or (b) which has been modified to  
10 improve its immunogenicity, wherein said modified polypeptide is at least 75% identical in amino acid sequence to the corresponding polypeptide of (a) or (b).

15 17. A fusion polypeptide comprising a polypeptide of claim 15 or 16 and an additional polypeptide.

18. The fusion polypeptide of claim 17 wherein the additional polypeptide is a heterologous signal peptide.

20

19. The fusion protein of claim 17 wherein the additional polypeptide has adjuvant activity.

20. A method for producing a polypeptide of claim 15 or  
25 16, comprising the step of culturing a unicellular host according to claim 12.

21. An antibody against the polypeptide of any one of claims 15 to 19.

5 22. A vaccine comprising at least one first polypeptide according to any one of claims 15 to 19 and a pharmaceutically acceptable carrier, optionally comprising a second polypeptide which enhances the immune response to the first polypeptide.

10 23. The vaccine of claim 22 wherein the second polypeptide comprises an additional *Chlamydia* polypeptide.

24. A pharmaceutical composition comprising a polypeptide according to any one of claims 15 to 19 and a pharmaceutically  
15 acceptable carrier.

25. A pharmaceutical composition comprising a vaccine according to claim 22 or 23 and a pharmaceutically acceptable carrier.

20

26. A pharmaceutical composition comprising an antibody according to claim 21 and a pharmaceutically acceptable carrier.

27. A method for preventing or treating *Chlamydia*  
25 infection using:

(a) the nucleic acid of any one of claims 1 to 7;

- (b) the vaccine of any one of claims 8, 9, 22 and 23;
- (c) the pharmaceutical composition of any one of claims 10, 11, 24 to 26;
- (d) the polypeptide of any one of claims 15 to 19; or
- 5 (e) the antibody of claim 21.

28. A method of detecting *Chlamydia* infection comprising the step of assaying a body fluid of a mammal to be tested, with a component selected from any one of:

- 10 (a) the nucleic acid of any one of claims 1 to 7;
- (b) the polypeptide of any one of claims 15 to 19; and
- (c) the antibody of claim 21.

29. A diagnostic kit comprising instructions for use and a  
15 component selected from any one of:

- (a) the nucleic acid of any one of claims 1 to 7;
- (b) the polypeptide of any one of claims 15 to 19; and
- (c) the antibody of claim 21.

20 30. A method for identifying a polypeptide of claims 15 to 19 which induces an immune response effective to prevent or lessen the severity of *Chlamydia* infection in a mammal previously immunized with polypeptide, comprising the steps of:

- (a) immunizing a mouse with the polypeptide; and
- 25 (b) inoculating the immunized mouse with *Chlamydia*;

wherein the polypeptide which prevents or lessens the severity of *Chlamydia* infection in the immunized mouse compared to a non-immunized control mouse is identified.

1/77

Figure 1: Nucleotide and deduced amino acid sequence of CPN100634

```

gataaaaatt cttgacagct gttttgtcat ctttaacttg atttacttat tttgtttcta 60
tattgatgcg aatagttctc taaaaaacaa aagcattacc atg aag act tcg att 115
                                         Met Lys Thr Ser Ile
                                         1       5

cct tgg gtt tta gtt tcc tcc gtg tta gct ttc tca tgt cac cta cag 163
Pro Trp Val Leu Val Ser Ser Val Leu Ala Phe Ser Cys His Leu Gln
          10              15              20

tca cta gct aac gag gaa ctt tta tta cct gat gat agc ttt aat gga 211
Ser Leu Ala Asn Glu Glu Leu Leu Ser Pro Asp Asp Ser Phe Asn Gly
          25              30              35

aat atc gat tca gga acg ttt act cca aaa act tca gcc aca aca tat 259
Asn Ile Asp Ser Gly Thr Phe Thr Pro Lys Thr Ser Ala Thr Thr Tyr
          40              45              50

tct cta aca gga gat gtc ttc ttt tac gag cct gga aaa ggc act ccc 307
Ser Leu Thr Gly Asp Val Phe Phe Tyr Glu Pro Gly Lys Gly Thr Pro
          55              60              65

tta tct gac agt tgt ttt aag caa acc acg gac aat ctt acc ttc ttg 355
Leu Ser Asp Ser Cys Phe Lys Gln Thr Thr Asp Asn Leu Thr Phe Leu
          70              75              80              85

ggg aac ggt cat agc tta acg ttt ggc ttt ata gat gct ggc act cat 403
Gly Asn Gly His Ser Leu Thr Phe Gly Phe Ile Asp Ala Gly Thr His
          90              95              100

gca ggt gct gct gca tct aca aca gca aat aag aat ctt acc ttc tca 451
Ala Gly Ala Ala Ala Ser Thr Thr Ala Asn Lys Asn Leu Thr Phe Ser
          105              110              115

ggg ttt tcc tta ctg agt ttt gat tcc tct cct agc aca acg gtt act 499
Gly Phe Ser Leu Leu Ser Phe Asp Ser Ser Pro Ser Thr Thr Val Thr
          120              125              130

aca ggt cag gga acg ctt tcc tca gca gga ggc gta aat tta gaa aat 547
Thr Gly Gln Gly Thr Leu Ser Ser Ala Gly Gly Val Asn Leu Glu Asn
          135              140              145

att cgt aaa ctt gta gtt gct ggg aat ttt tct act gca gat ggt gga 595
Ile Arg Lys Leu Val Val Ala Gly Asn Phe Ser Thr Ala Asp Gly Gly
          150              155              160              165

gct atc aaa gga gcg tct ttc ctt tta act ggc act tct gga gat gct 643
Ala Ile Lys Gly Ala Ser Phe Leu Leu Thr Gly Thr Ser Gly Asp Ala
          170              175              180

ctt ttt agt aac aac tct tca tca aca aag gga gga gca att gct act 691
Leu Phe Ser Asn Asn Ser Ser Ser Thr Lys Gly Gly Ala Ile Ala Thr
          185              190              195

```



2/77

Fig. 1 (con't)

|   |      |
|---|------|
| aca gca ggc gct cgc ata gca aat aac aca ggt tat gtt aga ttc cta | 739  |
| Thr Ala Gly Ala Arg Ile Ala Asn Asn Thr Gly Tyr Val Arg Phe Leu |      |
| 200 205 210   |      |
| tct aac ata gcg tct acg tca gga ggc gct atc gat gat gaa ggc acg | 787  |
| Ser Asn Ile Ala Ser Thr Ser Gly Gly Ala Ile Asp Asp Glu Gly Thr |      |
| 215 220 225   |      |
| tcg ata cta tcg aac aac aaa ttt cta tat ttt gaa ggg aat gca gcg | 835  |
| Ser Ile Leu Ser Asn Asn Lys Phe Leu Tyr Phe Glu Gly Asn Ala Ala |      |
| 230 235 240 245   |      |
| aaa act act ggc ggt gcg atc tgc aac acc aag gcg agt gga tct cct | 883  |
| Lys Thr Thr Gly Gly Ala Ile Cys Asn Thr Lys Ala Ser Gly Ser Pro |      |
| 250 255 260   |      |
| gaa ctg ata atc tct aac aat aag act ctg atc ttt gct tca aac gta | 931  |
| Glu Leu Ile Ile Ser Asn Asn Lys Thr Leu Ile Phe Ala Ser Asn Val |      |
| 265 270 275   |      |
| gca gaa aca agc ggt ggc gcc atc cat gct aaa aag cta gcc ctt tcc | 979  |
| Ala Glu Thr Ser Gly Gly Ala Ile His Ala Lys Lys Leu Ala Leu Ser |      |
| 280 285 290   |      |
| tct gga ggc ttt aca gag ttt cta cga aat aat gtc tca tca gca act | 1027 |
| Ser Gly Gly Phe Thr Glu Phe Leu Arg Asn Asn Val Ser Ser Ala Thr |      |
| 295 300 305   |      |
| cct aag ggg ggt gct atc agc atc gat gcc tca gga gag ctc agt ctt | 1075 |
| Pro Lys Gly Gly Ala Ile Ser Ile Asp Ala Ser Gly Glu Leu Ser Leu |      |
| 310 315 320 325   |      |
| tct gca gag aca gga aac att acc ttt gta aga aat acc ctt aca aca | 1123 |
| Ser Ala Glu Thr Gly Asn Ile Thr Phe Val Arg Asn Thr Leu Thr Thr |      |
| 330 335 340   |      |
| acc gga agt acc gat act cct aaa cgt aat gcg atc aac ata gga agt | 1171 |
| Thr Gly Ser Thr Asp Thr Pro Lys Arg Asn Ala Ile Asn Ile Gly Ser |      |
| 345 350 355   |      |
| aac ggg aaa ttc acg gaa tta cgg gct gct aaa aat cat aca att ttc | 1219 |
| Asn Gly Lys Phe Thr Glu Leu Arg Ala Ala Lys Asn His Thr Ile Phe |      |
| 360 365 370   |      |
| ttc tat gat ccc atc act tca gaa gga acc tca tca gac gta ttg aag | 1267 |
| Phe Tyr Asp Pro Ile Thr Ser Glu Gly Thr Ser Ser Asp Val Leu Lys |      |
| 375 380 385   |      |
| ata aat aac ggc tct gcg gga gct ctc aat cca tat caa gga acg att | 1315 |
| Ile Asn Asn Gly Ser Ala Gly Ala Leu Asn Pro Tyr Gln Gly Thr Ile |      |
| 390 395 400 405   |      |
| cta ttt tct gga gaa acc cta aca gca gat gaa ctt aaa gtt gct gac | 1363 |
| Leu Phe Ser Gly Glu Thr Leu Thr Ala Asp Glu Leu Lys Val Ala Asp |      |
| 410 415 420   |      |

3/77

Fig. 1 (con't)

|   |      |
|---|------|
| aat tta aaa tct tca ttc acg cag cca gtc tcc cta tcc gga gga aag | 1411 |
| Asn Leu Lys Ser Ser Phe Thr Gln Pro Val Ser Leu Ser Gly Gly Lys |      |
| 425 430 435   |      |
| tta ttg cta caa aag gga gtc act tta gag agc acg agc ttc tct caa | 1459 |
| Leu Leu Leu Gln Lys Gly Val Thr Leu Glu Ser Thr Ser Phe Ser Gln |      |
| 440 445 450   |      |
| gag gcc ggt tct ctc ctc ggc atg gat tca gga acg aca tta tca act | 1507 |
| Glu Ala Gly Ser Leu Leu Gly Met Asp Ser Gly Thr Thr Leu Ser Thr |      |
| 455 460 465   |      |
| aca gct ggg agt att aca atc acg aac cta gga atc aat gtt gac tcc | 1555 |
| Thr Ala Gly Ser Ile Thr Ile Thr Asn Leu Gly Ile Asn Val Asp Ser |      |
| 470 475 480 485   |      |
| tta ggt ctt aag cag ccc gtc agc cta aca gca aaa ggt gct tca aat | 1603 |
| Leu Gly Leu Lys Gln Pro Val Ser Leu Thr Ala Lys Gly Ala Ser Asn |      |
| 490 495 500   |      |
| aaa gtg atc gta tct ggg aag ctc aac ctg att gat att gaa ggg aac | 1651 |
| Lys Val Ile Val Ser Gly Lys Leu Asn Leu Ile Asp Ile Glu Gly Asn |      |
| 505 510 515   |      |
| att tat gaa agt cat atg ttc agc cat gac cag ctc ttc tct cta tta | 1699 |
| Ile Tyr Glu Ser His Met Phe Ser His Asp Gln Leu Phe Ser Leu Leu |      |
| 520 525 530   |      |
| aaa atc acg gtt gat gct gat gtt gat act aac gtt gac atc agc agc | 1747 |
| Lys Ile Thr Val Asp Ala Asp Val Asp Thr Asn Val Asp Ile Ser Ser |      |
| 535 540 545   |      |
| ctt atc cct gtt cct gct gag gat cct aat tca gaa tac gga ttc caa | 1795 |
| Leu Ile Pro Val Pro Ala Glu Asp Pro Asn Ser Glu Tyr Gly Phe Gln |      |
| 550 555 560 565   |      |
| gga caa tgg aat gtt aat tgg act acg gat aca gct aca aat aca aaa | 1843 |
| Gly Gln Trp Asn Val Asn Trp Thr Thr Asp Thr Ala Thr Asn Thr Lys |      |
| 570 575 580   |      |
| gag gcc acg gca act tgg acc aaa aca gga ttt gtt ccc agc ccc gaa | 1891 |
| Glu Ala Thr Ala Thr Trp Thr Lys Thr Gly Phe Val Pro Ser Pro Glu |      |
| 585 590 595   |      |
| aga aaa tct gcg tta gta tgc aat acc cta tgg gga gtc ttt act gac | 1939 |
| Arg Lys Ser Ala Leu Val Cys Asn Thr Leu Trp Gly Val Phe Thr Asp |      |
| 600 605 610   |      |
| att cgc tct ctg caa cag ctt gta gag atc ggc gca act ggt atg gaa | 1987 |
| Ile Arg Ser Leu Gln Gln Leu Val Glu Ile Gly Ala Thr Gly Met Glu |      |
| 615 620 625   |      |
| cac aaa caa ggt ttc tgg gtt tcc tcc atg acg aac ttc ctg cat aag | 2035 |
| His Lys Gln Gly Phe Trp Val Ser Ser Met Thr Asn Phe Leu His Lys |      |
| 630 635 640 645   |      |

4/77

Fig. 1 (con't)

|   |      |
|---|------|
| act gga gat gaa aat cgc aaa ggc ttc cgt cat acc tct gga ggc tac | 2083 |
| Thr Gly Asp Glu Asn Arg Lys Gly Phe Arg His Thr Ser Gly Gly Tyr |      |
| 650 655 660   |      |
| gtc atc ggt gga agt gct cac act cct aaa gac gac cta ttt acc ttt | 2131 |
| Val Ile Gly Gly Ser Ala His Thr Pro Lys Asp Asp Leu Phe Thr Phe |      |
| 665 670 675   |      |
| gcg ttc tgc cat ctc ttt gct aga gac aaa gat tgt ttt atc gct cac | 2179 |
| Ala Phe Cys His Leu Phe Ala Arg Asp Lys Asp Cys Phe Ile Ala His |      |
| 680 685 690   |      |
| aac aac tct aga acc tac ggt gga act tta ttc ttc aag cac tct cat | 2227 |
| Asn Asn Ser Arg Thr Tyr Gly Gly Thr Leu Phe Phe Lys His Ser His |      |
| 695 700 705   |      |
| acc cta caa ccc caa aac tat ttg aga tta gga aga gca aag ttt tct | 2275 |
| Thr Leu Gln Pro Gln Asn Tyr Leu Arg Leu Gly Arg Ala Lys Phe Ser |      |
| 710 715 720 725   |      |
| gaa tca gct ata gaa aaa ttc cct agg gaa att ccc cta gcc ttg gat | 2323 |
| Glu Ser Ala Ile Glu Lys Phe Pro Arg Glu Ile Pro Leu Ala Leu Asp |      |
| 730 735 740   |      |
| gtc caa gtt tcg ttc agc cat tca gac aac cgt atg gaa acg cac tat | 2371 |
| Val Gln Val Ser Phe Ser His Ser Asp Asn Arg Met Glu Thr His Tyr |      |
| 745 750 755   |      |
| acc tca ttg cca gaa tcc gaa ggt tct tgg agc aac gag tgt ata gct | 2419 |
| Thr Ser Leu Pro Glu Ser Glu Gly Ser Trp Ser Asn Glu Cys Ile Ala |      |
| 760 765 770   |      |
| ggg ggt atc ggc cta gac ctt cct ttt gtt ctt tcc aac cca cat cct | 2467 |
| Gly Gly Ile Gly Leu Asp Leu Pro Phe Val Leu Ser Asn Pro His Pro |      |
| 775 780 785   |      |
| ctt ttc aag acc ttc att cca cag atg aaa gtc gaa atg gtt tat gta | 2515 |
| Leu Phe Lys Thr Phe Ile Pro Gln Met Lys Val Glu Met Val Tyr Val |      |
| 790 795 800 805   |      |
| tca caa aat agc ttc ttc gaa agc tct agt gat ggc cgt ggt ttt agt | 2563 |
| Ser Gln Asn Ser Phe Phe Glu Ser Ser Ser Asp Gly Arg Gly Phe Ser |      |
| 810 815 820   |      |
| att gga agg ctg ctt aac ctc tcg att cct gtg ggt gcg aaa ttc gtg | 2611 |
| Ile Gly Arg Leu Leu Asn Leu Ser Ile Pro Val Gly Ala Lys Phe Val |      |
| 825 830 835   |      |
| cag ggg gat atc gga gat tcc tac acc tat gat ctc tca gga ttc ttt | 2659 |
| Gln Gly Asp Ile Gly Asp Ser Tyr Thr Tyr Asp Leu Ser Gly Phe Phe |      |
| 840 845 850   |      |
| gtt tcc gat gtc tat cgt aac aat ccc caa tct aca gcg act ctt grg | 2707 |
| Val Ser Asp Val Tyr Arg Asn Asn Pro Gln Ser Thr Ala Thr Leu Val |      |
| 855 860 865   |      |

5/77

Fig. 1 (con't)

```

atg agc cca gac tct tgg aaa att cgc ggt ggc aat ctt tca aga cag 2755
Met Ser Pro Asp Ser Trp Lys Ile Arg Gly Gly Asn Leu Ser Arg Gln
870                               875                               880                               885

gca ttt tta ctg agg ggt agc aac aac tac gtc tac aac tcc aat tgt 2803
Ala Phe Leu Leu Arg Gly Ser Asn Asn Tyr Val Tyr Asn Ser Asn Cys
                               890                               895                               900

gag ctc ttc gga cat tac gct atg gaa ctc cgt gga tct tca agg aac 2851
Glu Leu Phe Gly His Tyr Ala Met Glu Leu Arg Gly Ser Ser Arg Asn
                               905                               910                               915

tac aat gta gat gtt ggt acc aaa ctc cga ttc tagattgcta aaactcccta 2904
Tyr Asn Val Asp Val Gly Thr Lys Leu Arg Phe
                               920                               925

gttcttctag ggagttttct catactttta gggaaatatt tgctat 2950

```



7/77

Fig. 2 (con't)

```

                                BslI
                                EcoNI |
                                ScrFI | |
                                CviJI | | |
                                EcoRII | | |
                                | | |
                                BbsI
                                MboII
                                |
CviJI |
|
AACTTCAGCCACAACATATTCTCTAACAGGAGATGTCTTCTTTTACGAGCCTGAAAAGG
241 -----+-----+-----+-----+-----+ 300
TTGAAGTCGGTGTGTGATAAGAGATTGTCCTCTACAGAAGAAAATGCTCGGACCTTTTCC

                                CjePI  BsaJI
                                Hpy188IX TaaI  MseI| BstDSI
                                | | | |
                                | | | |
                                CACTCCCTTATCTGACAGTTGTTTAAAGCAAACCGGACAATCTTACCTTCTTGGGGAA
301 -----+-----+-----+-----+-----+ 360
GTGAGGGAATAGACTGTCAACAAAATTCGTTTGGTGCCTGTTAGAATGGAAGAACCCTT

                                CviRI
                                Fnu4HI |
                                TseI | |
                                Fnu4HI | | |
                                AlwNI | | |
                                BstAPI | | |
                                MwoI | | |
                                TseI | | |
                                AarI | | | |
                                MslI | | | |
                                CviRI | | | |
                                NlaIII | | | |
                                BbvI | | | | |
                                BbvI | | | | |
                                BspMI | | | | |
                                Cac8I | | | | |
                                | | | | |
                                CGGTCATAGCTTAAACGTTTGGCTTTATAGATGCTGGCACTCATGCAGGTGCTGCTGCATC
361 -----+-----+-----+-----+-----+ 420
GCCAGTATCGAATTGCAAACCGAAATATCTACGACCGTGAGTACGTCCACGACGACGTAG

                                BseMII
                                HinfI
                                BslI | DdeI HinfI
                                SfaNI TfiI DdeI | BseMII | TfiI
                                | | | |
                                TACAACAGCAAATAAGAATCTTACCTTCTCAGGGTTTTCTTACTGAGTTTTGATTCCCTC
421 -----+-----+-----+-----+-----+ 480
ATGTTGTCGTTTATTCTTAGAATGGAAGAGTCCCAAAGGAATGACTCAAACTAAGGAG

                                MnlI
                                MaeIII
                                MnlI |
                                BfaI | TaaI SfcI
                                | | |
                                TCCTAGCACAACGGTTACTACAGGTCAGGGAACGCTTTCCTCAGCAGGAGGCGTAAATTT
481 -----+-----+-----+-----+-----+ 540
AGGATCGTGTTGCCAATGATGTCCAGTCCCTTGCGAAAGGAGTCGTCTCCGCATTAA

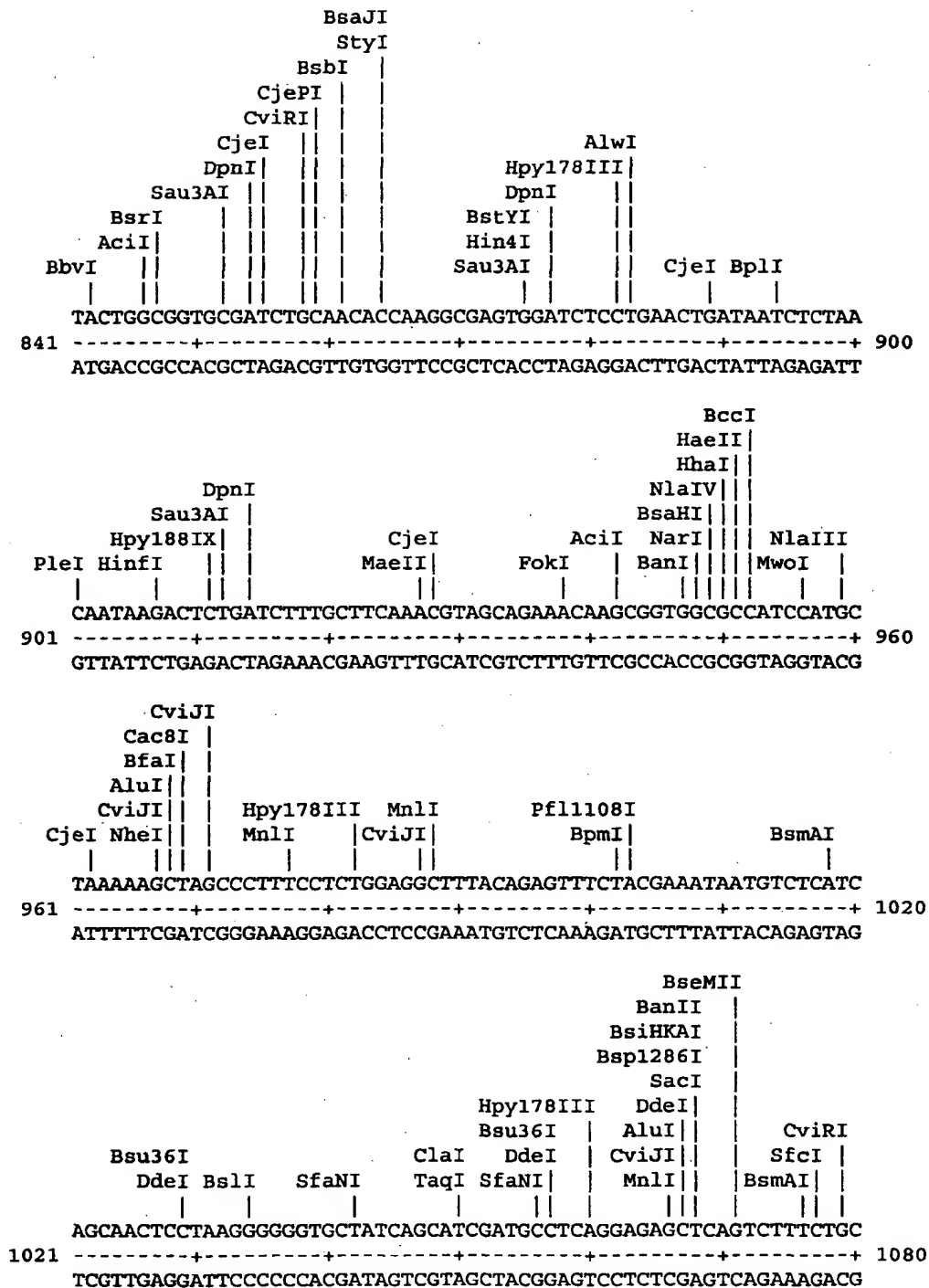
```

**Fig. 2 (con't)**



9/77

Fig. 2 (con't)





10/77

Fig. 2 (con't)

```

BseMII      MspI      BsaXI
PstI        BsaWI|    RsaI|
|           ||       |
AGAGACAGGAAACATTACCTTTGTAAGAAATACCCTTACAACAACCGGAAGTACCGATAC
1081 -----+-----+-----+-----+-----+-----+-----+ 1140
TCTCTGTCCTTTGTAATGGAACATTCTTTATGGGAATGTTGTTGGCCTTCATGGCTATG

                Tsp509I
                Sth132I|    Fnu4HI
                ApoI|      CviJI|
                DpnI      MaeIII Tsp509I|    TseI|
MaeII  Sau3AI| Sth132I| BscGI| BbvI| BscGI||
|      |      |      |      |      |      |
TCCTAAACGTAATGCGATCAACATAGGAAGTAACGGGAAATTACGGAATTACGGGCTGC
1141 -----+-----+-----+-----+-----+-----+-----+ 1200
AGGATTTCATTACGCTAGTTGTATCCTTCATTGCCCTTTAAGTGCCTTAATGCCCGACG

                DpnI      MnlI
                Sau3AI|    MaeII|
Tsp509I  Eco57I| | Hpy188IX Hpy188IX| |
MboII| AlwI| | BccI| NlaIV| |
|      |      |      |      |      |
TAAAAATCATACAATTTCTTCTATGATCCCATCACTTCAGAAGGAACCTCATCAGACGT
1201 -----+-----+-----+-----+-----+-----+-----+ 1260
ATTTTTAGTATGTTAAAGAAGATACTAGGGTAGTGAAGTCTTCCTTGGAGTAGTCTGCA

                BanII
                BceFI
                BsiHKA|
                CviJI  Bsp1286I
                FauI|    SacI
                MboII|    AluI|
                Sth132I| | AcI| CviJI|
                ||      |      |
ATTGAAGATAAATAACGGCTCTGCGGGAGCTCTCAATCCATATCAAGGAACGATTCTATT
1261 -----+-----+-----+-----+-----+-----+-----+ 1320
TAACCTTCTATTTATTGCCGAGACGCCCTCGAGAGTTAGGTATAGTTCCTTGCTAAGATAA

                DraI
                MseI|
                MboII||
Hpy178III      BpmI  MseI      Tsp509I| ||
|              |      |      |      |
TTCTGGAGAAACCCTAACAGCAGATGAACCTTAAAGTTGCTGACAATTTAAATCTTCATT
1321 -----+-----+-----+-----+-----+-----+-----+ 1380
AAGACCTCTTTGGGATTGTCGTCTACTTGAATTTCAACGACTGTTAAATTTTAGAAGTAA

```

11/77

Fig. 2 (con't)

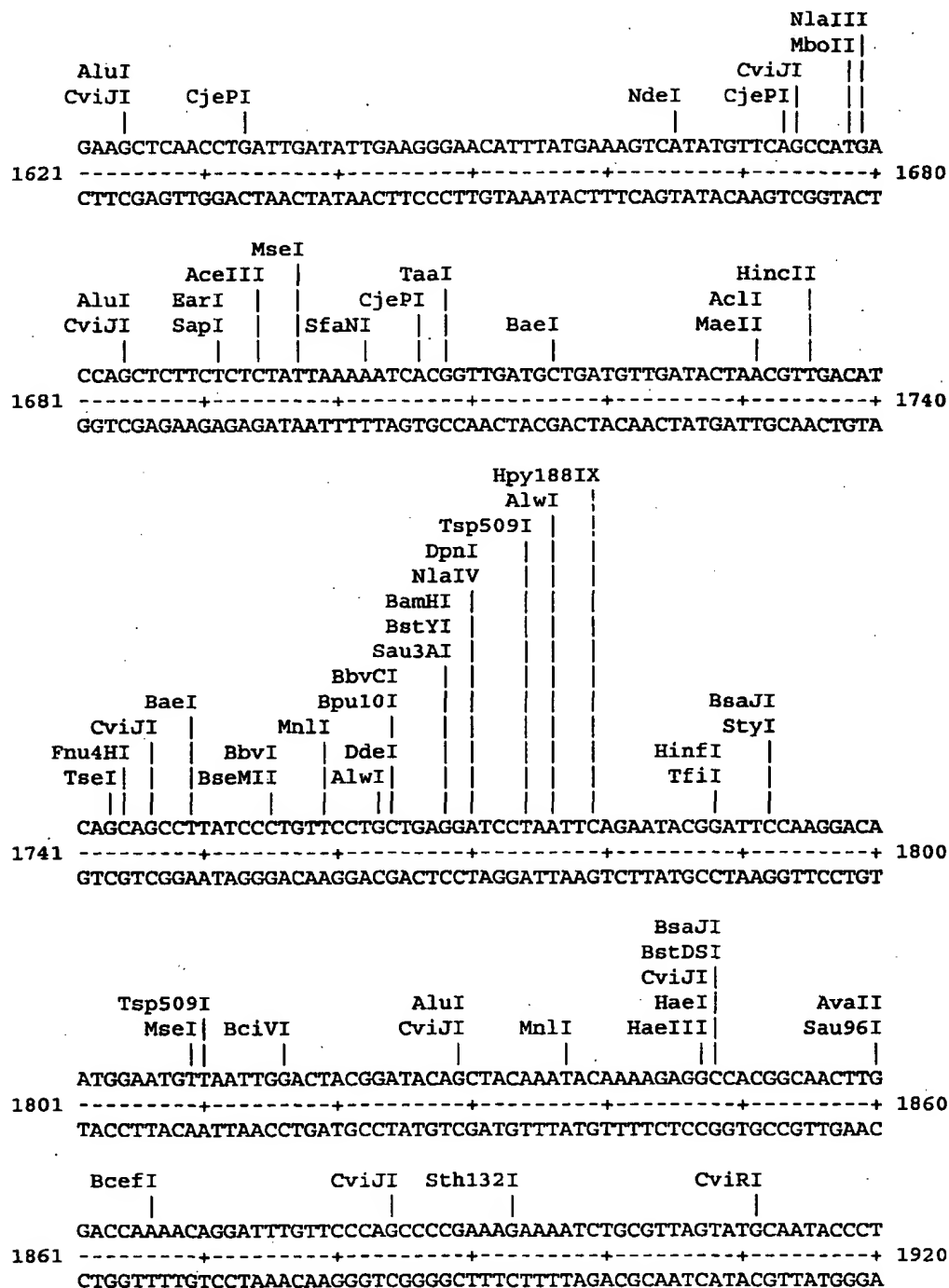
Hpy178III  
 MspII  
 BsaWI||  
 BspEI||  
 BsrI  
 CviJI| MnlI |||  
 Fnu4HI || BbvI |||  
 TseI| || BsmAI |||  
 || || ||  
 CACGCAGCCAGTCTCCCTATCCGGAGGAAAGTTATTGCTACAAAAGGGAGTCACTTTAGA  
 1381 -----+-----+-----+-----+-----+-----+-----+ 1440  
 GTGCGTCGGTCAGAGGGATAGGCCTCCTTTCAATAACGATGTTTTCCCTCAGTGAAATCT

MspI  
 BseRI|  
 BsrFI|  
 CviJI|  
 HaeIII|  
 Hpy178III ||  
 SmlI | ||  
 MnlI | ||  
 Hin4I | ||  
 AluI | ||  
 CviJI | ||  
 BsiHKA I || ||  
 Bsp1286I || ||  
 BssSI || ||  
 || ||  
 GAGCACGAGCTTCTCTCAAGAGGCCGGTTCTCTCCTCGGCATGGATTGAGGAACGACATT  
 1441 -----+-----+-----+-----+-----+-----+-----+ 1500  
 CTCGTGCTCGAAGAGAGTTCTCCGGCCAAGAGAGGAGCCGTACCTAAGTCCTTGCTGTAA

HinfI  
 AluI  
 CviJI  
 MspAI|  
 PvuII  
 SfcI |  
 |  
 ATCAACTACAGCTGGGAGTATTACAATCACGAACCTAGGAATCAATGTTGACTCCTTAGG  
 1501 -----+-----+-----+-----+-----+-----+-----+ 1560  
 TAGTTGATGTCGACCCTCATAATGTTAGTGCTTGGATCCTTAGTTACAACCTGAGGAATCC

BbvI  
 CviJI|  
 Sth132I|  
 BscGI ||  
 CviJI | ||  
 Fnu4HI || ||  
 MseI || ||  
 AflIII || ||  
 SmlI|TseI || ||  
 || || ||  
 TCTTAAGCAGCCCGTCAGCCTAACAGCAAAAGGTGCTTCAAATAAAGTGATCGTATCTGG  
 1561 -----+-----+-----+-----+-----+-----+-----+ 1620  
 AGAATTCGTCGGGCAGTCGGATTGTCGTTTTCCACGAAGTTTATTTCACTAGCATAGACC

**Fig. 2 (con't)**



13/77

Fig. 2 (con't)

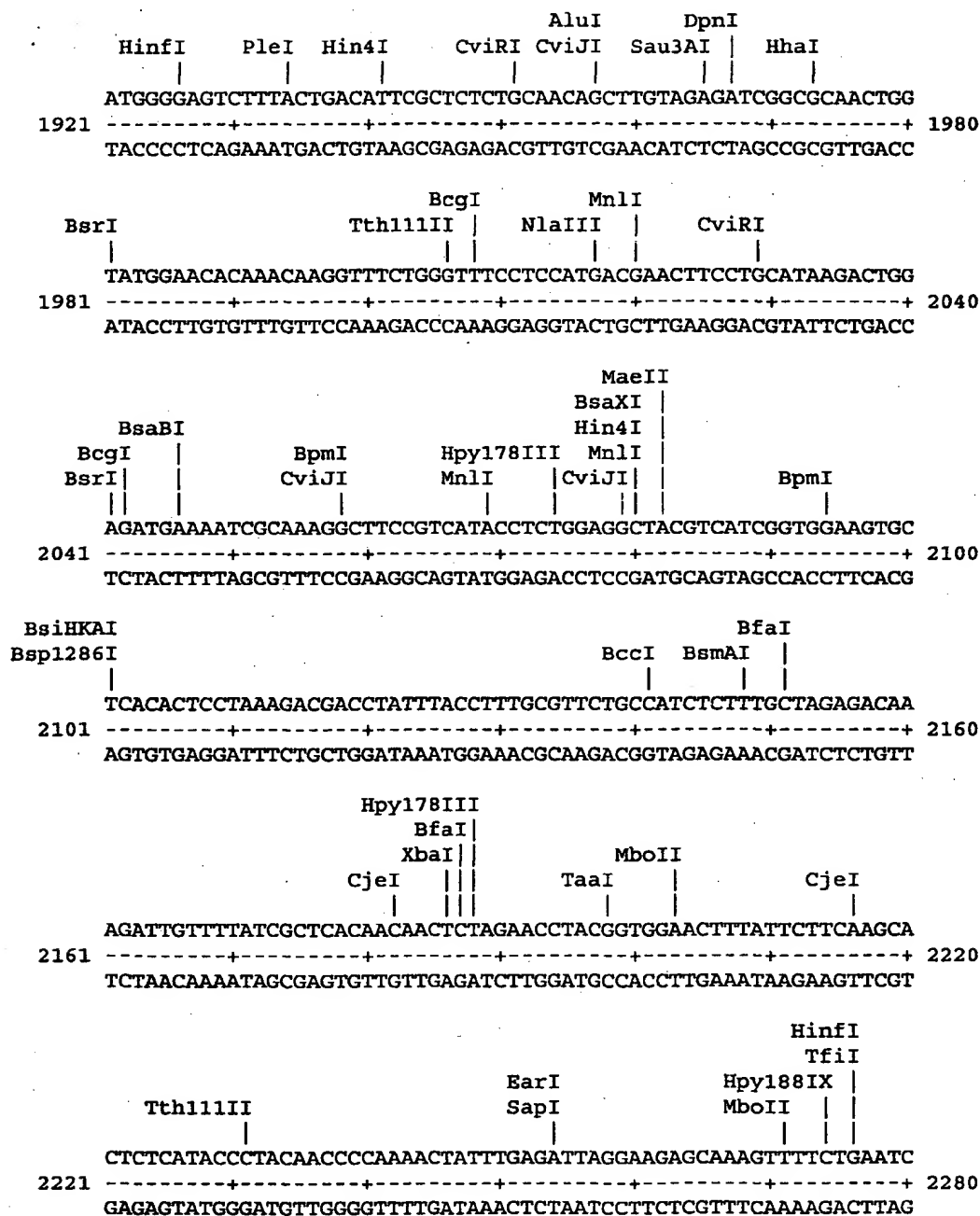
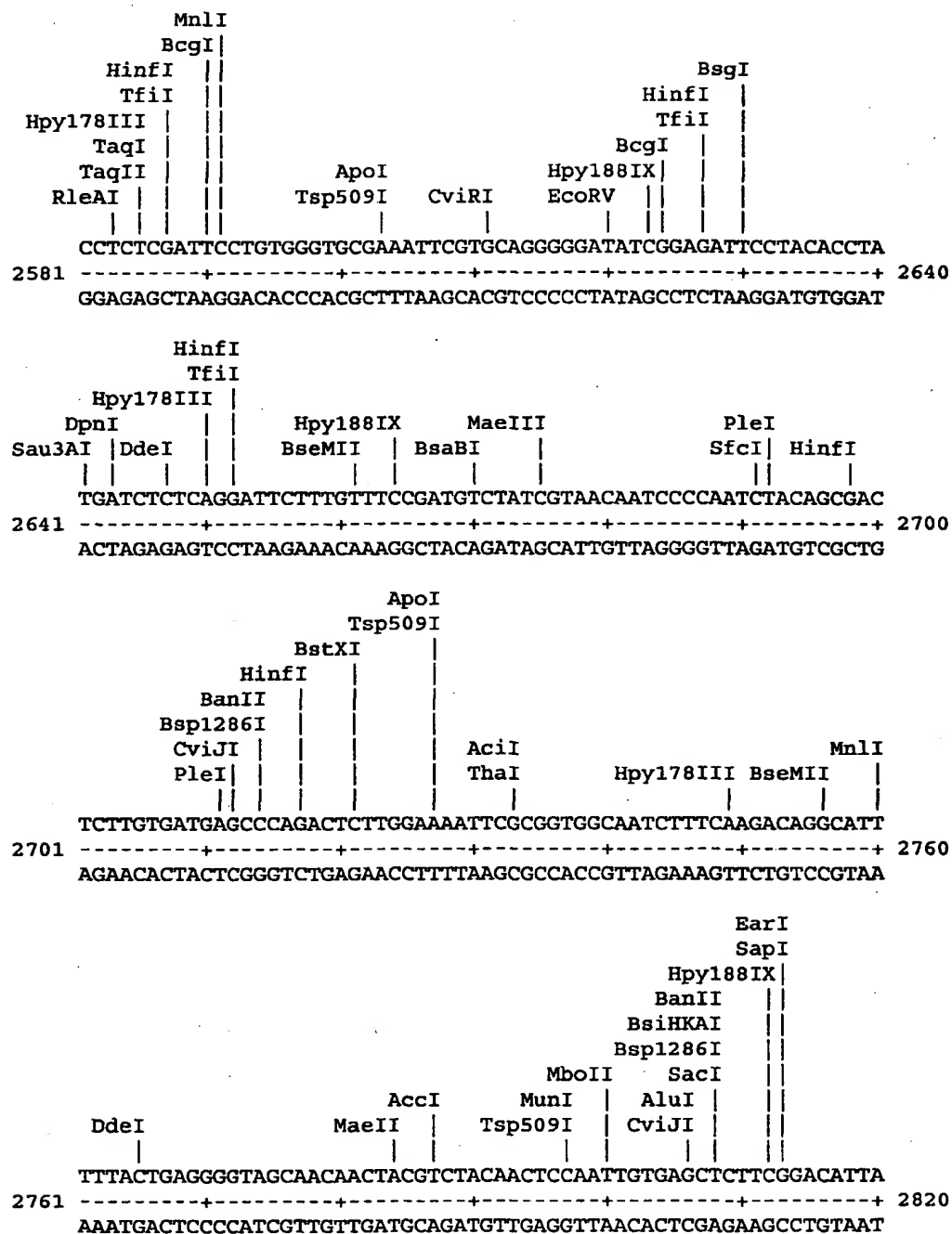


Fig. 2 (con't)



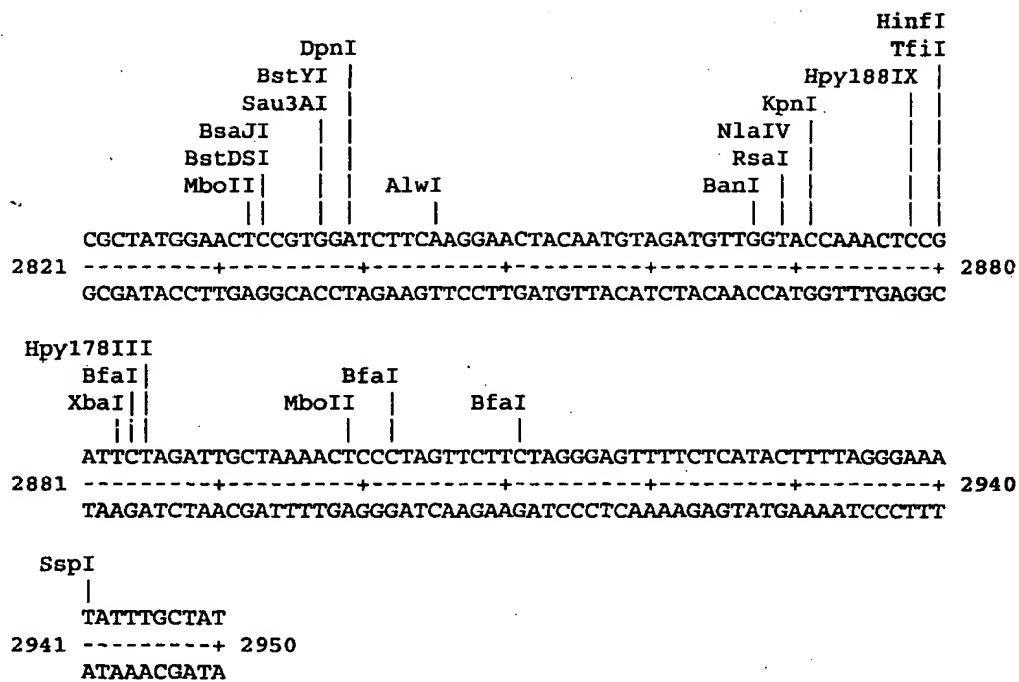
15/77

Fig. 2 (con't)



16/77

Fig. 2 (con't)



17/77

Figure 3: Nucleotide and deduced amino acid sequence of CPN100635

```

agttaaagat gacaaaacag ctgtcaagaa tttttatctt gactctctga gttttctatt 60
ttatatgacg caagtaagaa ttttaataata aagtggggtt atg aaa tcg caa ttt 115
                                         Met Lys Ser Gln Phe
                                         1       5

tcc tgg tta gtg ctc tct tcg aca ttg gca tgt ttt act agt tgt tcc 163
Ser Trp Leu Val Leu Ser Ser Thr Leu Ala Cys Phe Thr Ser Cys Ser
              10              15              20

act gtt ttt gct gca act gct gaa aat ata ggc ccc tct gat agc ttt 211
Thr Val Phe Ala Ala Thr Ala Glu Asn Ile Gly Pro Ser Asp Ser Phe
              25              30              35

gac gga agt act aac aca ggc acc tat act cct aaa aat acg act act 259
Asp Gly Ser Thr Asn Thr Gly Thr Tyr Thr Pro Lys Asn Thr Thr Thr
              40              45              50

gga ata gac tat act ctg aca gga gat ata act ctg caa aac ctt ggg 307
Gly Ile Asp Tyr Thr Leu Thr Gly Asp Ile Thr Leu Gln Asn Leu Gly
Gly Ile Asp Tyr Thr Leu Thr Gly Asp Ile Thr Leu Gln Asn Leu Gly
              55              60              65

gat tcg gca gct tta acg aag ggt tgt ttt tct gac act acg gaa tct 355
Asp Ser Ala Ala Leu Thr Lys Gly Cys Phe Ser Asp Thr Thr Glu Ser
Asp Ser Ala Ala Leu Thr Lys Gly Cys Phe Ser Asp Thr Thr Glu Ser
              70              75              80              85

tta agc ttt gcc ggt aag ggg tac tca ctt tct ttt tta aat att aag 403
Leu Ser Phe Ala Gly Lys Gly Tyr Ser Leu Ser Phe Leu Asn Ile Lys
Leu Ser Phe Ala Gly Lys Gly Tyr Ser Leu Ser Phe Leu Asn Ile Lys
              90              95              100

tct agt gct gaa ggc gca gcc ctt tct gtt aca act gat aaa aat ctg 451
Ser Ser Ala Glu Gly Ala Ala Leu Ser Val Thr Thr Asp Lys Asn Leu
Ser Ser Ala Glu Gly Ala Ala Leu Ser Val Thr Thr Asp Lys Asn Leu
              105              110              115

tcg cta aca gga ttt tcg agt ctt act ttc tta gcg gcc cca tca tcg 499
Ser Leu Thr Gly Phe Ser Ser Leu Thr Phe Leu Ala Ala Pro Ser Ser
Ser Leu Thr Gly Phe Ser Ser Leu Thr Phe Leu Ala Ala Pro Ser Ser
              120              125              130

gta atc aca acc ccc tca gga aaa ggt gca gtt aaa tgt gga ggg gat 547
Val Ile Thr Thr Pro Ser Gly Lys Gly Ala Val Lys Cys Gly Gly Asp
Val Ile Thr Thr Pro Ser Gly Lys Gly Ala Val Lys Cys Gly Gly Asp
              135              140              145

ctt aca ttt gat aac aat gga act att tta ttt aaa caa gat tac tgt 595
Leu Thr Phe Asp Asn Asn Gly Thr Ile Leu Phe Lys Gln Asp Tyr Cys
Leu Thr Phe Asp Asn Asn Gly Thr Ile Leu Phe Lys Gln Asp Tyr Cys
              150              155              160              165

```



18/77

Fig. 3 (con't)

|   |      |
|---|------|
| gag gaa aat ggc gga gcc att tct acc aag aat ctt tct ttg aaa aac | 643  |
| Glu Glu Asn Gly Gly Ala Ile Ser Thr Lys Asn Leu Ser Leu Lys Asn |      |
| Glu Glu Asn Gly Gly Ala Ile Ser Thr Lys Asn Leu Ser Leu Lys Asn |      |
| 170 175 180   |      |
| agc acg gga tcg att tct ttt gaa ggg aat aaa tcg agc gca aca ggg | 691  |
| Ser Thr Gly Ser Ile Ser Phe Glu Gly Asn Lys Ser Ser Ala Thr Gly |      |
| Ser Thr Gly Ser Ile Ser Phe Glu Gly Asn Lys Ser Ser Ala Thr Gly |      |
| 185 190 195   |      |
| aaa aaa ggt ggg gct att tgt gct act ggt act gta gat att aca aat | 739  |
| Lys Lys Gly Gly Ala Ile Cys Ala Thr Gly Thr Val Asp Ile Thr Asn |      |
| Lys Lys Gly Gly Ala Ile Cys Ala Thr Gly Thr Val Asp Ile Thr Asn |      |
| 200 205 210   |      |
| aat acg gct cct acc ctc ttc tcg aac aat att gct gaa gct gca ggt | 787  |
| Asn Thr Ala Pro Thr Leu Phe Ser Asn Asn Ile Ala Glu Ala Ala Gly |      |
| Asn Thr Ala Pro Thr Leu Phe Ser Asn Asn Ile Ala Glu Ala Ala Gly |      |
| 215 220 225   |      |
| gga gct ata aat agc aca gga aac tgt aca att aca ggg aat acg tct | 835  |
| Gly Ala Ile Asn Ser Thr Gly Asn Cys Thr Ile Thr Gly Asn Thr Ser |      |
| Gly Ala Ile Asn Ser Thr Gly Asn Cys Thr Ile Thr Gly Asn Thr Ser |      |
| 230 235 240 245   |      |
| ctt gta ttt tct gaa aat agt gtg aca gcg acc gca gga aat gga gga | 883  |
| Leu Val Phe Ser Glu Asn Ser Val Thr Ala Thr Ala Gly Asn Gly Gly |      |
| Leu Val Phe Ser Glu Asn Ser Val Thr Ala Thr Ala Gly Asn Gly Gly |      |
| 250 255 260   |      |
| gct ctt tct gga gat gcc gat gtt acc ata tct ggg aat cag agt gta | 931  |
| Ala Leu Ser Gly Asp Ala Asp Val Thr Ile Ser Gly Asn Gln Ser Val |      |
| Ala Leu Ser Gly Asp Ala Asp Val Thr Ile Ser Gly Asn Gln Ser Val |      |
| 265 270 275   |      |
| act ttc tca gga aac caa gct gta gct aat ggc gga gcc att tat gct | 979  |
| Thr Phe Ser Gly Asn Gln Ala Val Ala Asn Gly Gly Ala Ile Tyr Ala |      |
| Thr Phe Ser Gly Asn Gln Ala Val Ala Asn Gly Gly Ala Ile Tyr Ala |      |
| 280 285 290   |      |
| aag aag ctt aca ctg gct tcc ggg ggg ggg ggg ggg aat ccc ttt tct | 1027 |
| Lys Lys Leu Thr Leu Ala Ser Gly Gly Gly Gly Gly Asn Pro Phe Ser |      |
| Lys Lys Leu Thr Leu Ala Ser Gly Gly Gly Gly Gly Asn Pro Phe Ser |      |
| 295 300 305   |      |
| aac aat ata gtc caa ggt acc act gca ggt aat ggt gga gcc att tct | 1075 |
| Asn Asn Ile Val Gln Gly Thr Thr Ala Gly Asn Gly Gly Ala Ile Ser |      |
| Asn Asn Ile Val Gln Gly Thr Thr Ala Gly Asn Gly Gly Ala Ile Ser |      |
| 310 315 320 325   |      |
| ata ctg gca gct gga gag tgt agt ctt ttc agc gaa gca ggg gac cat | 1123 |
| Ile Leu Ala Ala Gly Glu Cys Ser Leu Phe Ser Glu Ala Gly Asp His |      |
| Ile Leu Ala Ala Gly Glu Cys Ser Leu Phe Ser Glu Ala Gly Asp His |      |
| 330 335 340   |      |

19/77

Fig. 3 (con't)

|   |      |
|---|------|
| tac ctt aat ggg aat gcc att gtt gca act aca cca caa act aca aaa | 1171 |
| Tyr Leu Asn Gly Asn Ala Ile Val Ala Thr Thr Pro Gln Thr Thr Lys |      |
| Tyr Leu Asn Gly Asn Ala Ile Val Ala Thr Thr Pro Gln Thr Thr Lys |      |
| 345 350 355   |      |
| aga aat tct att gac ata gga tct act ggc aaa gat cac gaa tta cgt | 1219 |
| Arg Asn Ser Ile Asp Ile Gly Ser Thr Gly Lys Asp His Glu Leu Arg |      |
| Arg Asn Ser Ile Asp Ile Gly Ser Thr Gly Lys Asp His Glu Leu Arg |      |
| 360 365 370   |      |
| gca ata tct ggg cat agc atc ttt ttc tac gat ccg att act gct aat | 1267 |
| Ala Ile Ser Gly His Ser Ile Phe Phe Tyr Asp Pro Ile Thr Ala Asn |      |
| Ala Ile Ser Gly His Ser Ile Phe Phe Tyr Asp Pro Ile Thr Ala Asn |      |
| 375 380 385   |      |
| acg gct gcg gat tct aca gat act tta aat ctc aat aag gct gat gca | 1315 |
| Thr Ala Ala Asp Ser Thr Asp Thr Leu Asn Leu Asn Lys Ala Asp Ala |      |
| Thr Ala Ala Asp Ser Thr Asp Thr Leu Asn Leu Asn Lys Ala Asp Ala |      |
| 390 395 400 405   |      |
| ggg aat agt aca gat tat agt ggg tcg att gtt ttt tct ggt gaa aag | 1363 |
| Gly Asn Ser Thr Asp Tyr Ser Gly Ser Ile Val Phe Ser Gly Glu Lys |      |
| Gly Asn Ser Thr Asp Tyr Ser Gly Ser Ile Val Phe Ser Gly Glu Lys |      |
| 410 415 420   |      |
| ctc tct gaa gat gaa gca aaa gtt gca gac aac ctc act tct acg ctg | 1411 |
| Leu Ser Glu Asp Glu Ala Lys Val Ala Asp Asn Leu Thr Ser Thr Leu |      |
| Leu Ser Glu Asp Glu Ala Lys Val Ala Asp Asn Leu Thr Ser Thr Leu |      |
| 425 430 435   |      |
| aag cag cct gta act cta act gca gga aat tta gta ctt aaa cgt ggt | 1459 |
| Lys Gln Pro Val Thr Leu Thr Ala Gly Asn Leu Val Leu Lys Arg Gly |      |
| Lys Gln Pro Val Thr Leu Thr Ala Gly Asn Leu Val Leu Lys Arg Gly |      |
| 440 445 450   |      |
| gtc act ctc gat acg aaa ggc ttt act cag acc gcg ggt tcc tct gtt | 1507 |
| Val Thr Leu Asp Thr Lys Gly Phe Thr Gln Thr Ala Gly Ser Ser Val |      |
| Val Thr Leu Asp Thr Lys Gly Phe Thr Gln Thr Ala Gly Ser Ser Val |      |
| 455 460 465   |      |
| att atg gat gcg ggc aca acg tta aaa gca agt aca gag gag gtc act | 1555 |
| Ile Met Asp Ala Gly Thr Thr Leu Lys Ala Ser Thr Glu Glu Val Thr |      |
| Ile Met Asp Ala Gly Thr Thr Leu Lys Ala Ser Thr Glu Glu Val Thr |      |
| 470 475 480 485   |      |
| tta aca ggt ctt tcc att cct gta gac tct tta ggc gag ggt aag aaa | 1603 |
| Leu Thr Gly Leu Ser Ile Pro Val Asp Ser Leu Gly Glu Gly Lys Lys |      |
| Leu Thr Gly Leu Ser Ile Pro Val Asp Ser Leu Gly Glu Gly Lys Lys |      |
| 490 495 500   |      |
| gtt gta att gct gct tct gca gca agt aaa aat gta gcc ctt agt ggt | 1651 |
| Val Val Ile Ala Ala Ser Ala Ala Ser Lys Asn Val Ala Leu Ser Gly |      |
| Val Val Ile Ala Ala Ser Ala Ala Ser Lys Asn Val Ala Leu Ser Gly |      |
| 505 510 515   |      |

20/77

Fig. 3 (con't)

|   |      |
|---|------|
| ccg att ctt ctt ttg gat aac caa ggg aat gct tat gaa aat cac gac | 1699 |
| Pro Ile Leu Leu Leu Asp Asn Gln Gly Asn Ala Tyr Glu Asn His Asp |      |
| Pro Ile Leu Leu Leu Asp Asn Gln Gly Asn Ala Tyr Glu Asn His Asp |      |
| 520 525 530   |      |
| tta gga aaa act caa gac ttt tca ttt gtg cag ctc tct gct ctg ggt | 1747 |
| Leu Gly Lys Thr Gln Asp Phe Ser Phe Val Gln Leu Ser Ala Leu Gly |      |
| Leu Gly Lys Thr Gln Asp Phe Ser Phe Val Gln Leu Ser Ala Leu Gly |      |
| 535 540 545   |      |
| act gca aca act aca gat gtt cca gcg gtt cct aca gta gca act cct | 1795 |
| Thr Ala Thr Thr Thr Asp Val Pro Ala Val Pro Thr Val Ala Thr Pro |      |
| Thr Ala Thr Thr Thr Asp Val Pro Ala Val Pro Thr Val Ala Thr Pro |      |
| 550 555 560 565   |      |
| acg cac tat ggg tat caa ggt act tgg gga atg act tgg gtt gat gat | 1843 |
| Thr His Tyr Gly Tyr Gln Gly Thr Trp Gly Met Thr Trp Val Asp Asp |      |
| Thr His Tyr Gly Tyr Gln Gly Thr Trp Gly Met Thr Trp Val Asp Asp |      |
| 570 575 580   |      |
| acc gca agc act cca aag act aag aca gcg aca tta gct tgg acc aat | 1891 |
| Thr Ala Ser Thr Pro Lys Thr Lys Thr Ala Thr Leu Ala Trp Thr Asn |      |
| Thr Ala Ser Thr Pro Lys Thr Lys Thr Ala Thr Leu Ala Trp Thr Asn |      |
| 585 590 595   |      |
| aca ggc tac ctt ccg aat cct gag cgt caa gga cct tta gtt cct aat | 1939 |
| Thr Gly Tyr Leu Pro Asn Pro Glu Arg Gln Gly Pro Leu Val Pro Asn |      |
| Thr Gly Tyr Leu Pro Asn Pro Glu Arg Gln Gly Pro Leu Val Pro Asn |      |
| 600 605 610   |      |
| agc ctt tgg gga tct ttt tca gac atc caa gcg att caa ggt gtc ata | 1987 |
| Ser Leu Trp Gly Ser Phe Ser Asp Ile Gln Ala Ile Gln Gly Val Ile |      |
| Ser Leu Trp Gly Ser Phe Ser Asp Ile Gln Ala Ile Gln Gly Val Ile |      |
| 615 620 625   |      |
| gag aga agt gct ttg act ctt tgt tca gat cga ggc ttc tgg gct gcg | 2035 |
| Glu Arg Ser Ala Leu Thr Leu Cys Ser Asp Arg Gly Phe Trp Ala Ala |      |
| Glu Arg Ser Ala Leu Thr Leu Cys Ser Asp Arg Gly Phe Trp Ala Ala |      |
| 630 635 640 645   |      |
| gga gtc gcc aat ttc tta gat aaa gat aag aaa ggg gaa aaa cgc aaa | 2083 |
| Gly Val Ala Asn Phe Leu Asp Lys Asp Lys Lys Gly Glu Lys Arg Lys |      |
| Gly Val Ala Asn Phe Leu Asp Lys Asp Lys Lys Gly Glu Lys Arg Lys |      |
| 650 655 660   |      |
| tac cgt cat aaa tct ggt gga tat gct atc gga ggt gca gcg caa act | 2131 |
| Tyr Arg His Lys Ser Gly Gly Tyr Ala Ile Gly Gly Ala Ala Gln Thr |      |
| Tyr Arg His Lys Ser Gly Gly Tyr Ala Ile Gly Gly Ala Ala Gln Thr |      |
| 665 670 675   |      |
| tgt tct gaa aac tta att agc ttt gcc ttt tgc caa ctc ttt ggt agc | 2179 |
| Cys Ser Glu Asn Leu Ile Ser Phe Ala Phe Cys Gln Leu Phe Gly Ser |      |
| Cys Ser Glu Asn Leu Ile Ser Phe Ala Phe Cys Gln Leu Phe Gly Ser |      |
| 680 685 690   |      |

21/77

Fig. 3 (con't)

|   |      |
|---|------|
| gat aaa gat ttc tta gtc gct aaa aat cat act gat acc tat gca gga | 2227 |
| Asp Lys Asp Phe Leu Val Ala Lys Asn His Thr Asp Thr Tyr Ala Gly |      |
| Asp Lys Asp Phe Leu Val Ala Lys Asn His Thr Asp Thr Tyr Ala Gly |      |
| 695 700 705   |      |
| gcc ttc tat atc caa cac att aca gaa tgt agt ggg ttc ata ggt tgt | 2275 |
| Ala Phe Tyr Ile Gln His Ile Thr Glu Cys Ser Gly Phe Ile Gly Cys |      |
| Ala Phe Tyr Ile Gln His Ile Thr Glu Cys Ser Gly Phe Ile Gly Cys |      |
| 710 715 720 725   |      |
| ctc tta gat aaa ctt cct ggc tct tgg agt cat aaa ccc ctc gtt tta | 2323 |
| Leu Leu Asp Lys Leu Pro Gly Ser Trp Ser His Lys Pro Leu Val Leu |      |
| Leu Leu Asp Lys Leu Pro Gly Ser Trp Ser His Lys Pro Leu Val Leu |      |
| 730 735 740   |      |
| gaa ggg cag ctc gct tat agc cac gtc agt aat gat ctg aag aca aag | 2371 |
| Glu Gly Gln Leu Ala Tyr Ser His Val Ser Asn Asp Leu Lys Thr Lys |      |
| Glu Gly Gln Leu Ala Tyr Ser His Val Ser Asn Asp Leu Lys Thr Lys |      |
| 745 750 755   |      |
| tat act gcg tat cct gag gtg aaa ggt tct tgg ggg aat aat gct ttt | 2419 |
| Tyr Thr Ala Tyr Pro Glu Val Lys Gly Ser Trp Gly Asn Asn Ala Phe |      |
| Tyr Thr Ala Tyr Pro Glu Val Lys Gly Ser Trp Gly Asn Asn Ala Phe |      |
| 760 765 770   |      |
| aac atg atg ttg gga gct tct tct cat tct tat cct gaa tac ctg cat | 2467 |
| Asn Met Met Leu Gly Ala Ser Ser His Ser Tyr Pro Glu Tyr Leu His |      |
| Asn Met Met Leu Gly Ala Ser Ser His Ser Tyr Pro Glu Tyr Leu His |      |
| 775 780 785   |      |
| tgt ttt gat acc tat gct cca tac atc aaa ctg aat ctg acc tat ata | 2515 |
| Cys Phe Asp Thr Tyr Ala Pro Tyr Ile Lys Leu Asn Leu Thr Tyr Ile |      |
| Cys Phe Asp Thr Tyr Ala Pro Tyr Ile Lys Leu Asn Leu Thr Tyr Ile |      |
| 790 795 800 805   |      |
| cgt cag gac agc ttc tcg gag aaa ggt aca gaa gga aga tct ttt gat | 2563 |
| Arg Gln Asp Ser Phe Ser Glu Lys Gly Thr Glu Gly Arg Ser Phe Asp |      |
| Arg Gln Asp Ser Phe Ser Glu Lys Gly Thr Glu Gly Arg Ser Phe Asp |      |
| 810 815 820   |      |
| gac agc aac ctc ttc aat tta tct ttg cct ata ggg gtg aag ttt gag | 2611 |
| Asp Ser Asn Leu Phe Asn Leu Ser Leu Pro Ile Gly Val Lys Phe Glu |      |
| Asp Ser Asn Leu Phe Asn Leu Ser Leu Pro Ile Gly Val Lys Phe Glu |      |
| 825 830 835   |      |
| aag ttc tct gat tgt aat gac ttt tct tat gat ctg act tta tcc tat | 2659 |
| Lys Phe Ser Asp Cys Asn Asp Phe Ser Tyr Asp Leu Thr Leu Ser Tyr |      |
| Lys Phe Ser Asp Cys Asn Asp Phe Ser Tyr Asp Leu Thr Leu Ser Tyr |      |
| 840 845 850   |      |
| gtt cct gat ctt atc cgc aat gat ccc aaa tgc act aca gca ctt gta | 2707 |
| Val Pro Asp Leu Ile Arg Asn Asp Pro Lys Cys Thr Thr Ala Leu Val |      |
| Val Pro Asp Leu Ile Arg Asn Asp Pro Lys Cys Thr Thr Ala Leu Val |      |
| 855 860 865   |      |

22/77

Fig. 3 (con't)

```

atc agc gga gcc tct tgg gaa act tat gcc aat aac tta gca cga cag 2755
Ile Ser Gly Ala Ser Trp Glu Thr Tyr Ala Asn Asn Leu Ala Arg Gln
Ile Ser Gly Ala Ser Trp Glu Thr Tyr Ala Asn Asn Leu Ala Arg Gln
870                      875                      880                      885

gcc ttg caa gtg cgt gca ggc agt cac tac gcc ttc tct cct atg ttt 2803
Ala Leu Gln Val Arg Ala Gly Ser His Tyr Ala Phe Ser Pro Met Phe
Ala Leu Gln Val Arg Ala Gly Ser His Tyr Ala Phe Ser Pro Met Phe
                        890                      895                      900

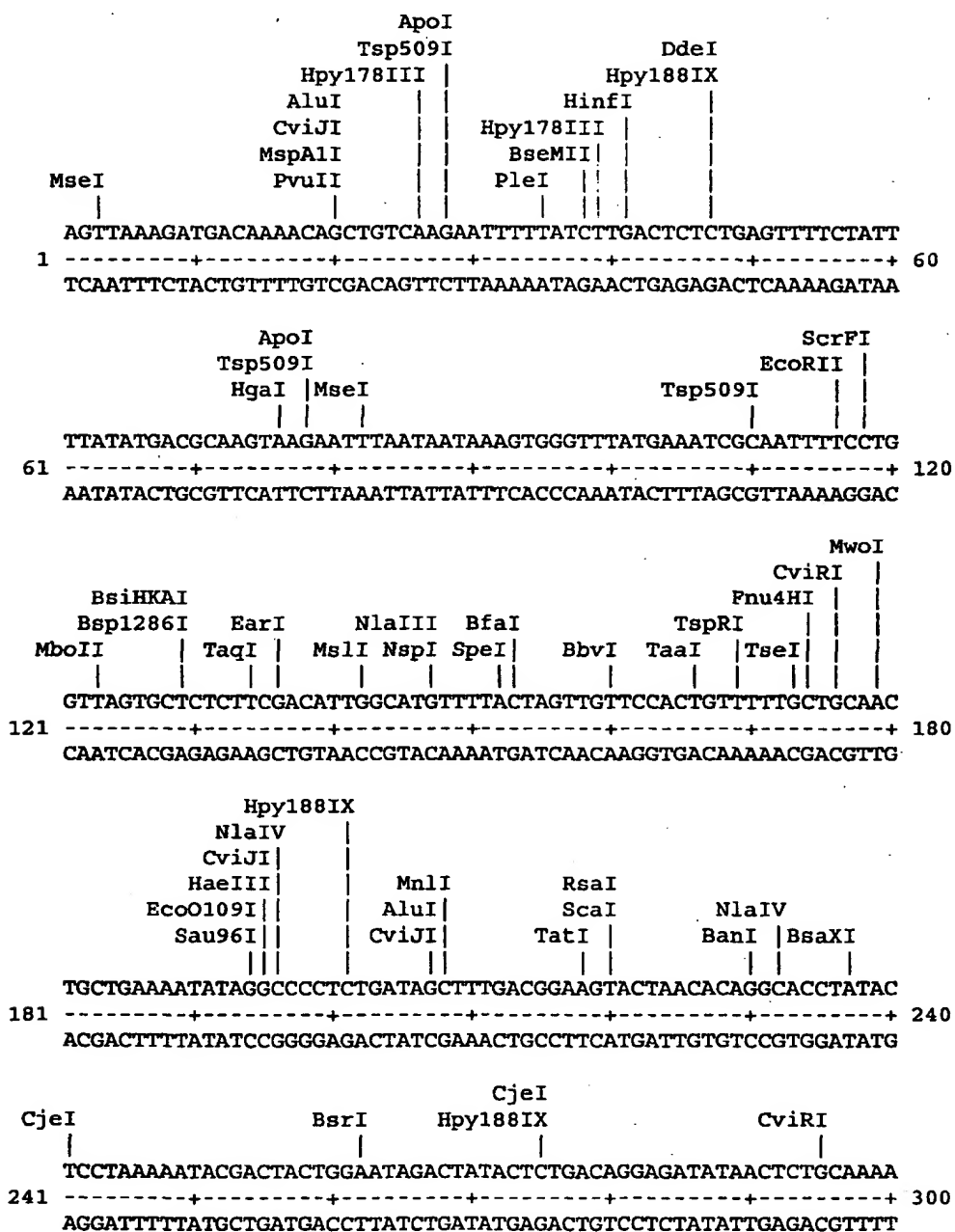
gaa gtg ctc ggc cag ttt gtc ttt gaa gtt cgt gga tcc tca cgg att 2851
Glu Val Leu Gly Gln Phe Val Phe Glu Val Arg Gly Ser Ser Arg Ile
Glu Val Leu Gly Gln Phe Val Phe Glu Val Arg Gly Ser Ser Arg Ile
                        905                      910                      915

tat aat gta gat ctt ggg ggt aag ttc caa ttc taggagcgtc tctcatgtct 2904
Tyr Asn Val Asp Leu Gly Gly Lys Phe Gln Phe
Tyr Asn Val Asp Leu Gly Gly Lys Phe Gln Phe
                        920                      925

cagaaattct gagagagatc gcatttagga ttttcttaaa cacgac 2950

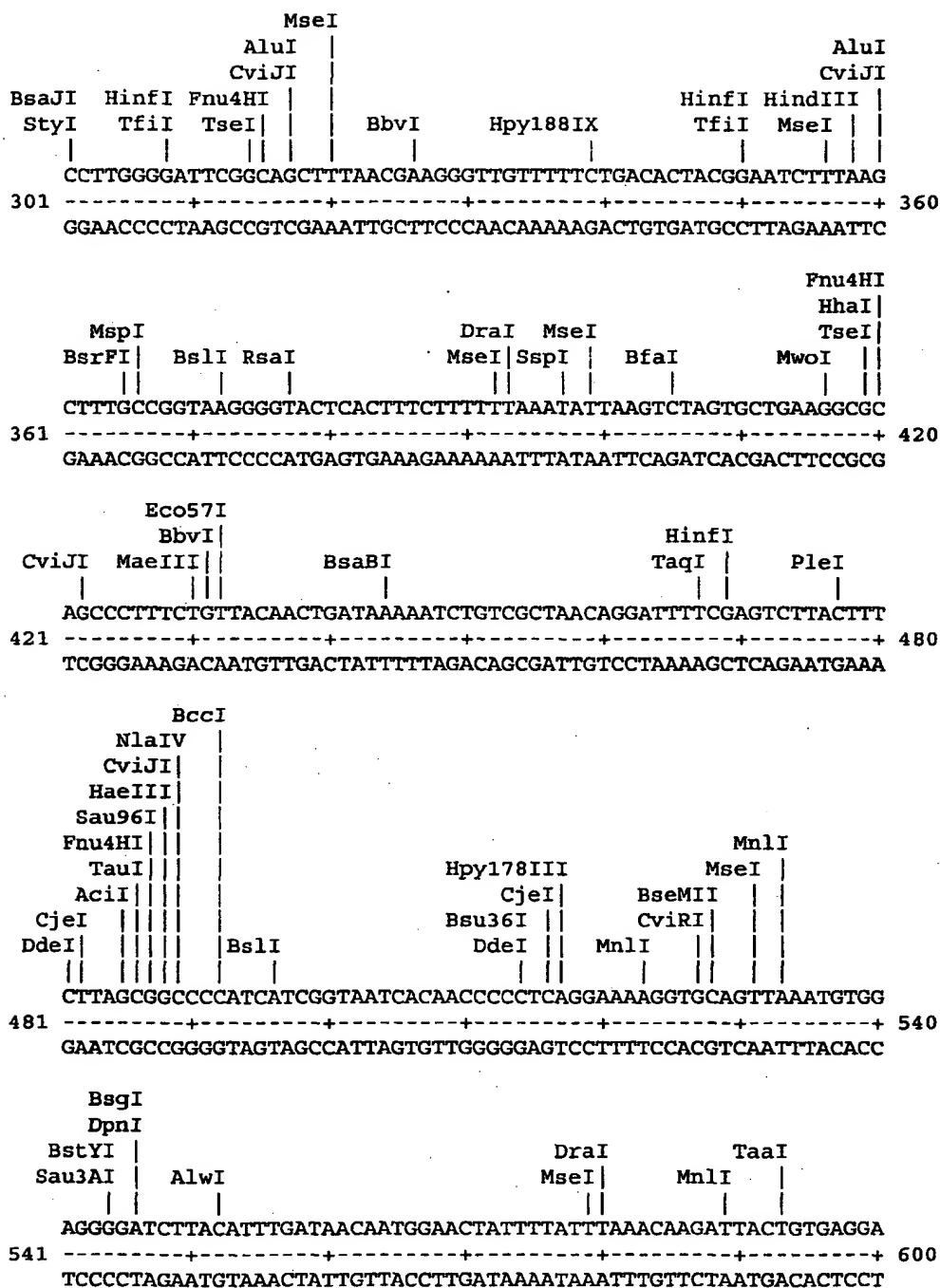
```

**Figure 4: Restriction enzyme analysis of CPN100635**



24/77

Fig. 4 (con't)



**Fig. 4 (con't)**





**Fig. 4 (con't)**

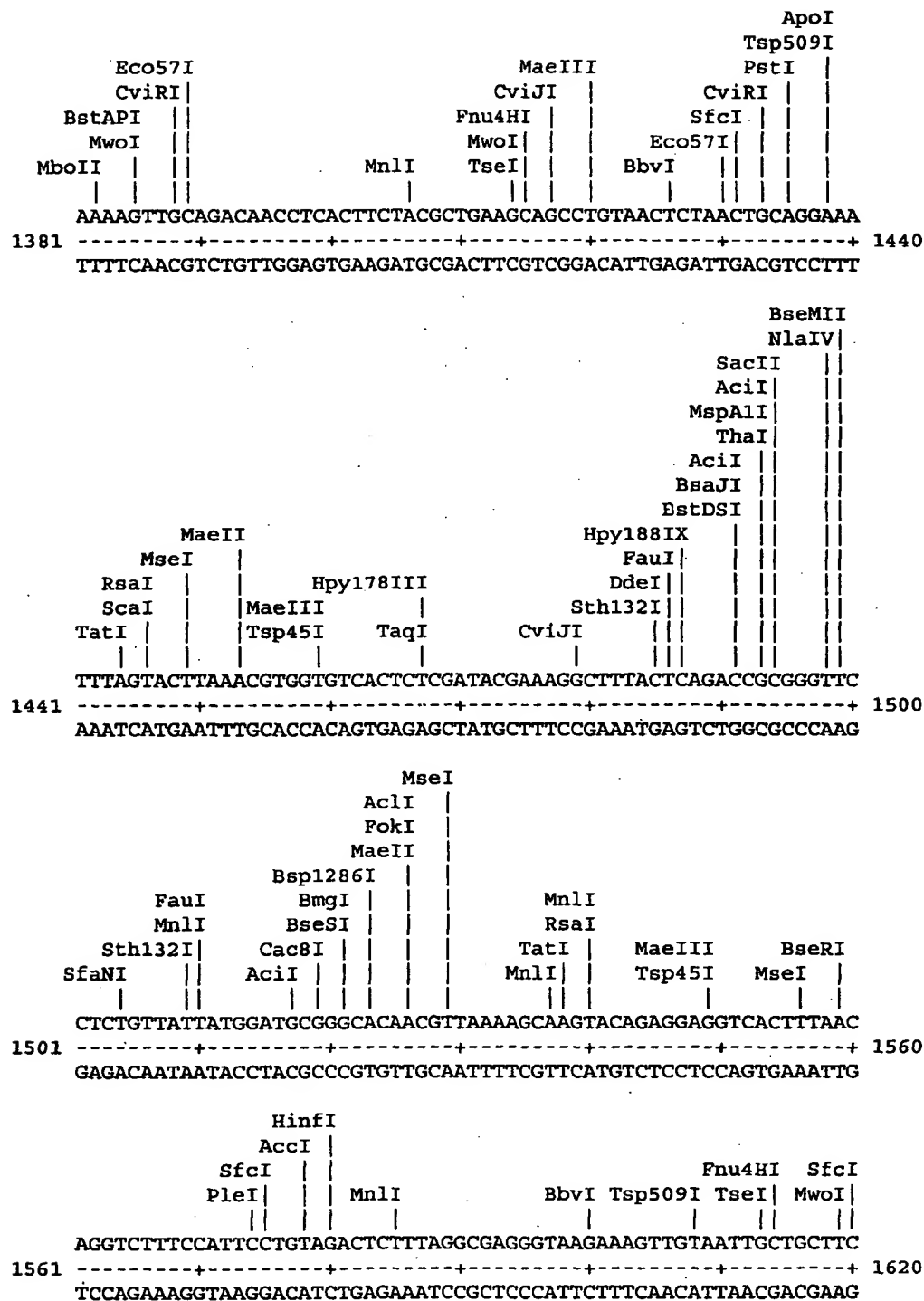


**Fig. 4 (con't)**



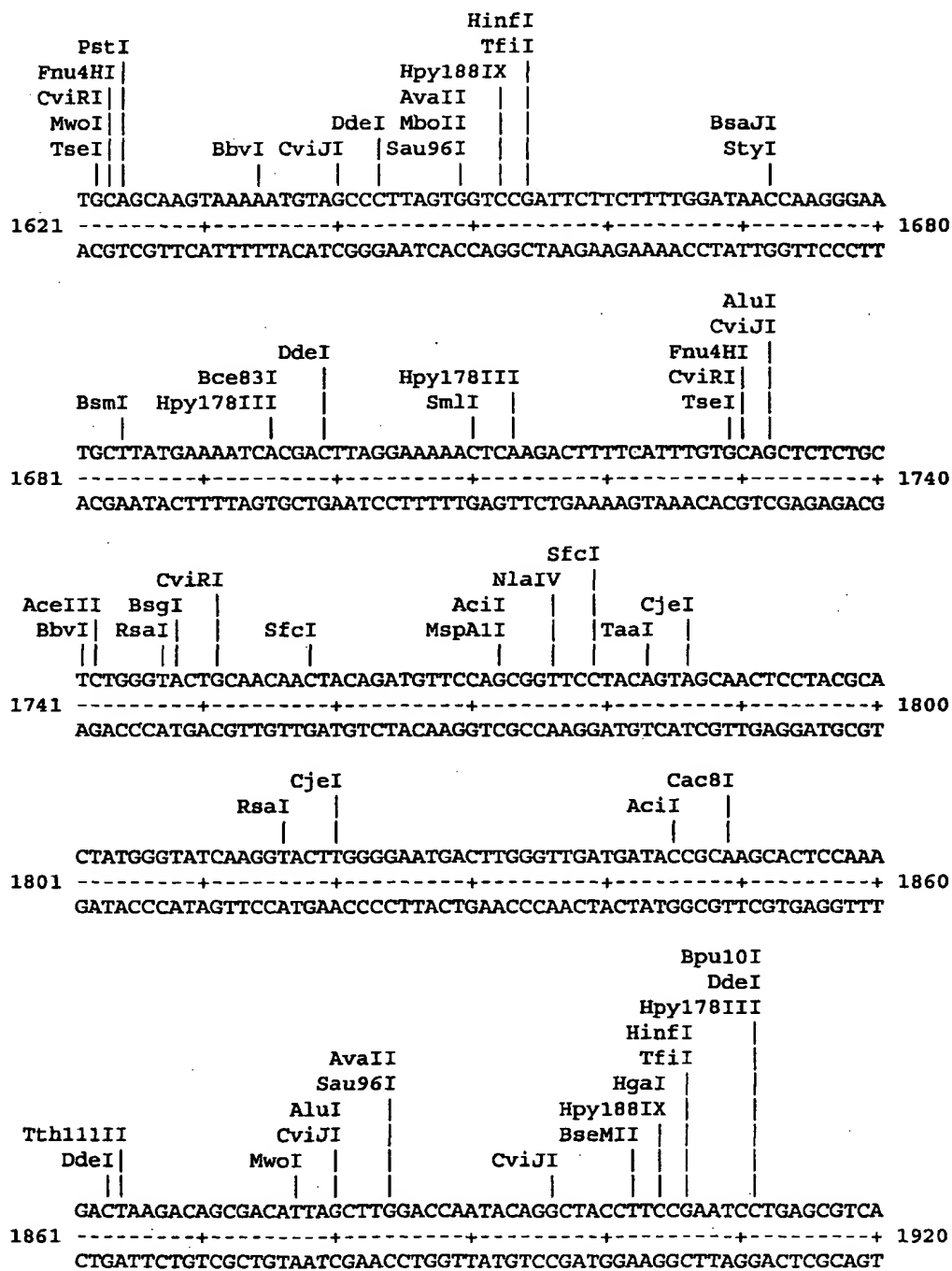
28/77

Fig. 4 (con't)



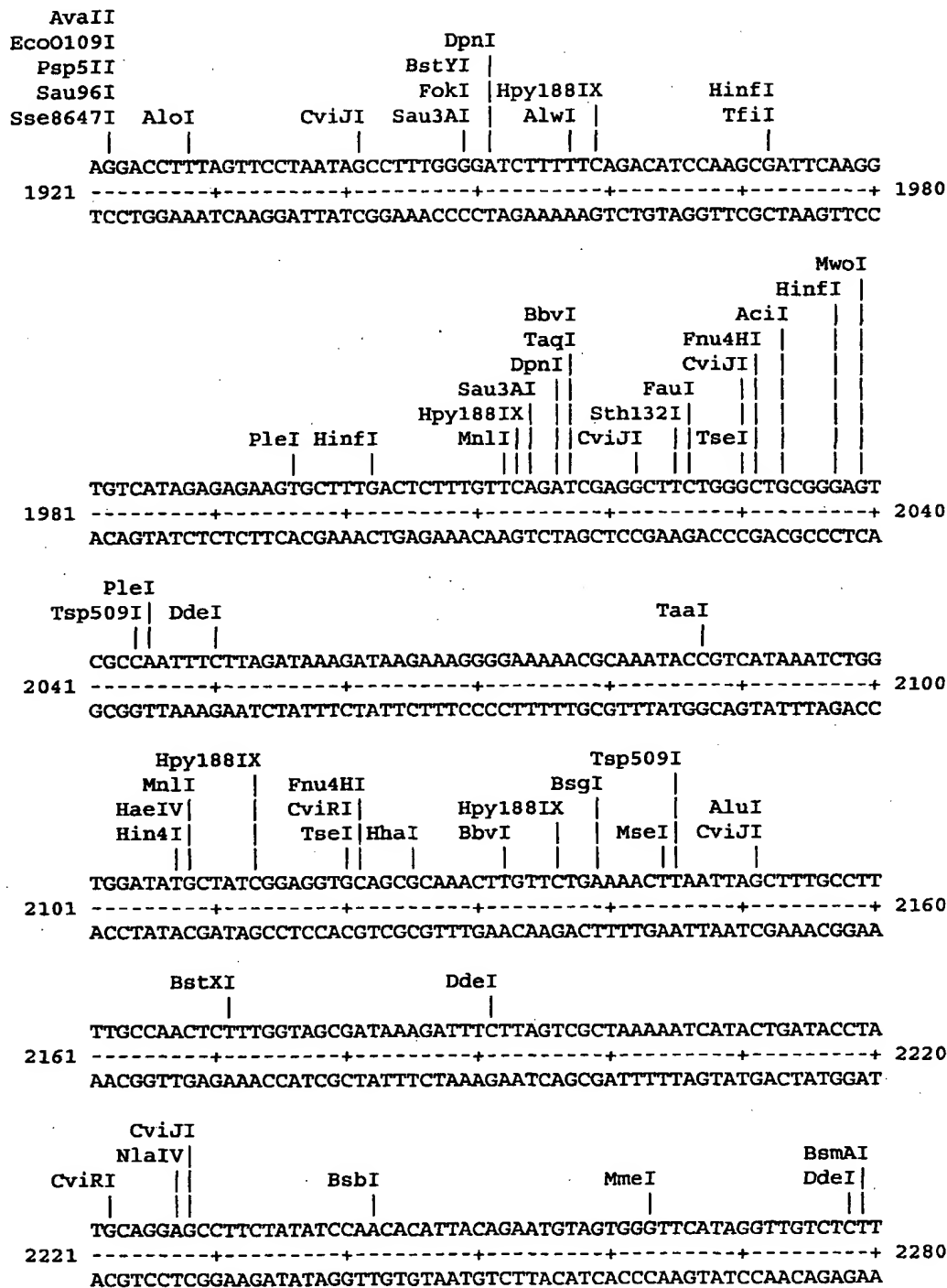
29/77

Fig. 4 (con't)



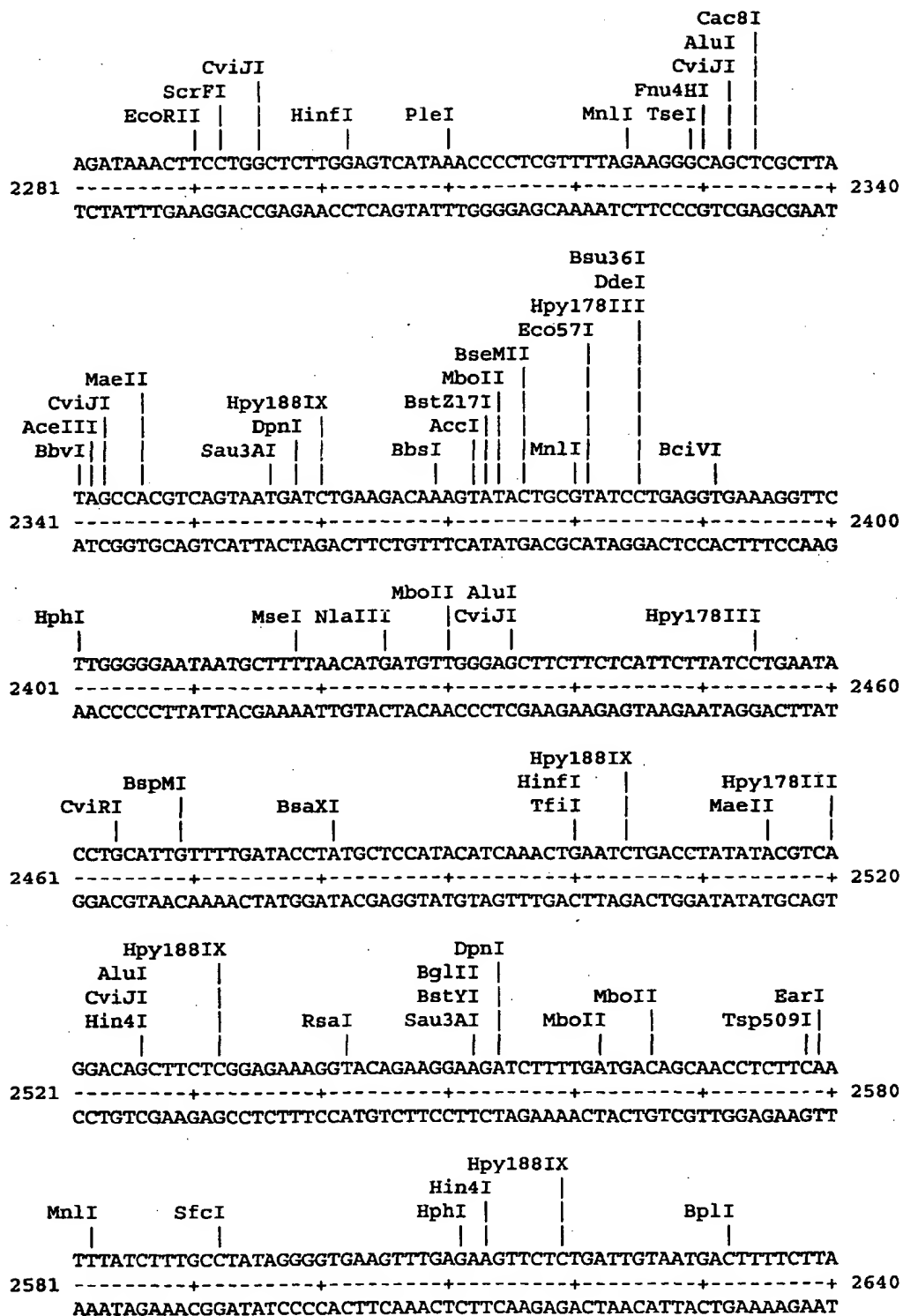
30/77

Fig. 4 (con't)



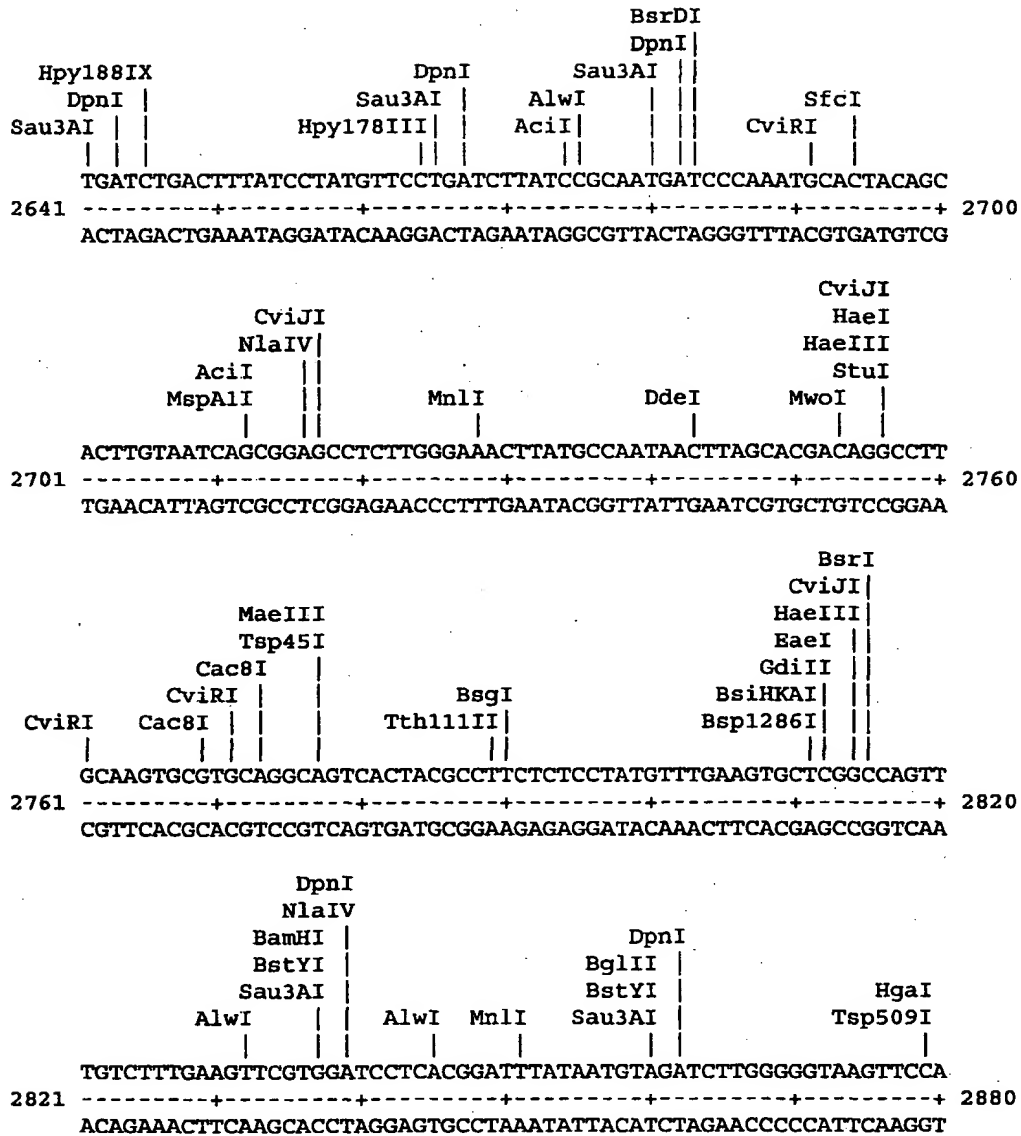
31/77

Fig. 4 (con't)



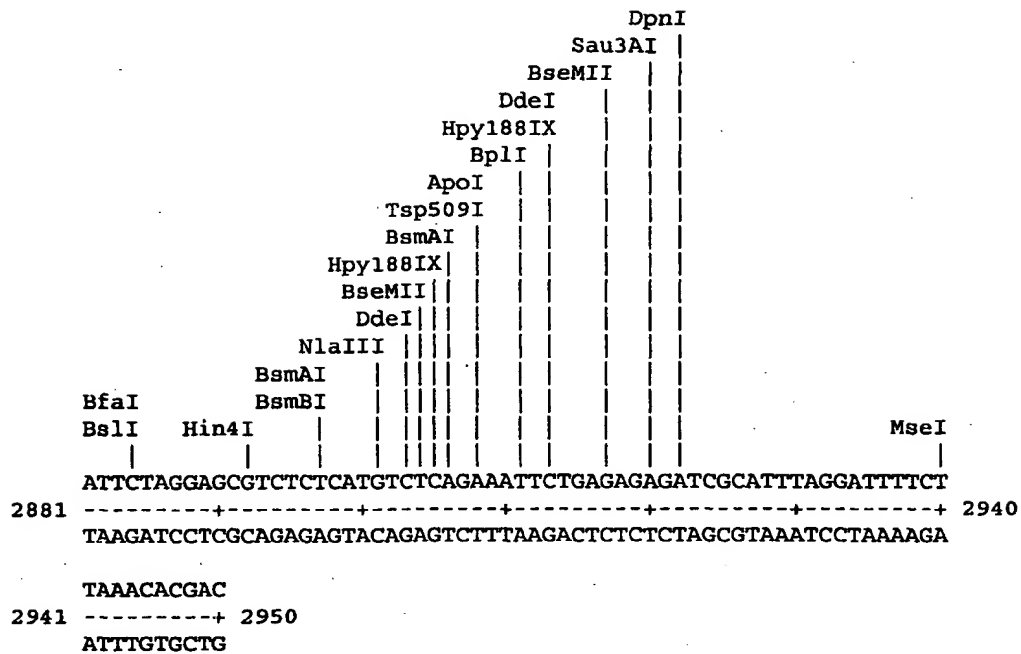
32/77

Fig. 4 (con't)



33/77

Fig. 4 (con't)





34/77

Figure 5: Nucleotide and deduced amino acid sequence of CPN100638

```

tgtagattct taacttacta gtctctcctt tcctcttgct ttctttaatt tattgcagta 60
tgtggtgaaa taatttggtta aaccacctat agccctctac atg aaa tcc tct ctt 115
                                         Met Lys Ser Ser Leu
                                         1       5
cat tgg ttt tta atc tcg tca tct tta gca ctt ccc ttg tca cta aat 163
His Trp Phe Leu Ile Ser Ser Ser Leu Ala Leu Pro Leu Ser Leu Asn
          10              15              20
ttc tct gcg ttt gct gct gtt gtt gaa atc aat cta gga cct acc aat 211
Phe Ser Ala Phe Ala Ala Val Val Glu Ile Asn Leu Gly Pro Thr Asn
          25              30              35
agc ttc tct gga cca gga acc tac act cct cca gcc caa aca aca aat 259
Ser Phe Ser Ser Gly Pro Gly Thr Tyr Thr Pro Pro Ala Gln Thr Thr Asn
          40              45              50
gca gat gga act atc tat aat cta aca ggg gat gtc tca atc acc aat 307
Ala Asp Gly Thr Ile Tyr Asn Leu Thr Gly Asp Val Ser Ile Thr Asn
          55              60              65
gca gga tct ccg aca gct cta acc gct tcc tgc ttt aaa gaa act act 355
Ala Gly Ser Pro Thr Ala Leu Thr Ala Ser Cys Phe Lys Glu Thr Thr
          70              75              80              85
ggg aat ctt tct ttc caa ggc cac ggc tac caa ttt ctc cta caa aat 403
Gly Asn Leu Ser Phe Gln Gly His Gly Tyr Gln Phe Leu Leu Gln Asn
          90              95              100
atc gat gcg gga gcg aac tgt acc ttt acc aat aca gct gca aat aag 451
Ile Asp Ala Gly Ala Asn Cys Thr Phe Thr Asn Thr Ala Ala Asn Lys
          105              110              115
ctt ctc tcc ttt tca gga ttc tcc tat ttg tca cta ata caa acc acg 499
Leu Leu Ser Phe Ser Gly Phe Ser Tyr Leu Ser Leu Ile Gln Thr Thr
          120              125              130
aat gct acc aca gga aca gga gcc atc aag tcc aca gga gct tgt tct 547
Asn Ala Thr Thr Gly Thr Gly Ala Ile Lys Ser Thr Gly Ala Cys Ser
          135              140              145
att cag tcg aac tat agt tgc tac ttt ggc caa aac ttt tct aat gac 595
Ile Gln Ser Asn Tyr Ser Cys Tyr Phe Gly Gln Asn Phe Ser Asn Asp
          150              155              160              165
aat gga gcc gcc ctc caa gcc agc tct atc agt cta tcg cta aac ccc 643
Asn Gly Gly Ala Leu Gln Gly Ser Ser Ile Ser Leu Ser Leu Asn Pro
          170              175              180
aac cta acg ttt gcc aaa aac aaa gca acg caa aaa ggg ggt gcc ctc 691
Asn Leu Thr Phe Ala Lys Asn Lys Ala Thr Gln Lys Gly Gly Ala Leu
          185              190              195

```

SUBSTITUTE SHEET (RULE 26)

35/77

Fig. 5 (con't)

|   |      |
|---|------|
| tat tcc acg gga ggg att aca att aac aat acg tta aac tca gca tca | 739  |
| Tyr Ser Thr Gly Gly Ile Thr Ile Asn Asn Thr Leu Asn Ser Ala Ser |      |
| 200 205 210   |      |
| ttt tct gaa aat acc gcg gcg aac aat ggc gga gcc att tac acg gaa | 787  |
| Phe Ser Glu Asn Thr Ala Ala Asn Asn Gly Gly Ala Ile Tyr Thr Glu |      |
| 215 220 225   |      |
| gct agc agt ttt att agc agc aac aaa gca att agc ttt ata aac aat | 835  |
| Ala Ser Ser Phe Ile Ser Ser Asn Lys Ala Ile Ser Phe Ile Asn Asn |      |
| 230 235 240 245   |      |
| agt gtg acc gca acc tca gct aca ggg gga gcc att tac tgt agt agt | 883  |
| Ser Val Thr Ala Thr Ser Ala Thr Gly Gly Ala Ile Tyr Cys Ser Ser |      |
| 250 255 260   |      |
| aca tca gcc ccc aaa cca gtc tta act cta tca gac aac ggg gaa ctg | 931  |
| Thr Ser Ala Pro Lys Pro Val Leu Thr Leu Ser Asp Asn Gly Glu Leu |      |
| 265 270 275   |      |
| aac ttt ata gga aat aca gca att act agt ggt ggg gcg att tat act | 979  |
| Asn Phe Ile Gly Asn Thr Ala Ile Thr Ser Gly Gly Ala Ile Tyr Thr |      |
| 280 285 290   |      |
| gac aat cta gtt ctt tct tct gga gga cct acg ctt ttt aaa aac aac | 1027 |
| Asp Asn Leu Val Leu Ser Ser Gly Gly Pro Thr Leu Phe Lys Asn Asn |      |
| 295 300 305   |      |
| tct ggc tat gat act gca gct ccc tta gga gga gca att gcg att gct | 1075 |
| Ser Gly Tyr Asp Thr Ala Ala Pro Leu Gly Gly Ala Ile Ala Ile Ala |      |
| 310 315 320 325   |      |
| gac tct gga tct ttg agt ctt tcg gct ctt ggt gga gac atc act ttt | 1123 |
| Asp Ser Gly Ser Leu Ser Leu Ser Ala Leu Gly Gly Asp Ile Thr Phe |      |
| 330 335 340   |      |
| gaa gga aac aca gta gtc aaa gga gct tct tcg agt cag acc act acc | 1171 |
| Glu Gly Asn Thr Val Val Lys Gly Ala Ser Ser Ser Gln Thr Thr Thr |      |
| 345 350 355   |      |
| aga aat tct att aac atc gga aac acc aat gct aag att gta cag ctg | 1219 |
| Arg Asn Ser Ile Asn Ile Gly Asn Thr Asn Ala Lys Ile Val Gln Leu |      |
| 360 365 370   |      |
| cga gcc tct caa ggc aat act atc tac ttc tat gat cct ata aca act | 1267 |
| Arg Ala Ser Gln Gly Asn Thr Ile Tyr Phe Tyr Asp Pro Ile Thr Thr |      |
| 375 380 385   |      |
| agc atc act gca gct ctc tca gat gct cta aac tta aat ggt cct gac | 1315 |
| Ser Ile Thr Ala Ala Leu Ser Asp Ala Leu Asn Leu Asn Gly Pro Asp |      |
| 390 395 400 405   |      |
| ctt gca ggg aat cct gca tat caa gga acc atc gta ttt tct gga gag | 1363 |
| Leu Ala Gly Asn Pro Ala Tyr Gln Gly Thr Ile Val Phe Ser Gly Glu |      |
| 410 415 420   |      |

36/77

Fig. 5 (con't)

|   |      |
|---|------|
| aag ctc tcg gaa gca gaa gct gca gaa gct gat aat ctc aaa tct aca | 1411 |
| Lys Leu Ser Glu Ala Glu Ala Ala Glu Ala Asp Asn Leu Lys Ser Thr |      |
| 425 430 435   |      |
| att cag caa cct cta act ctt gcg gga ggg caa ctc tct ctt aaa tca | 1459 |
| Ile Gln Gln Pro Leu Thr Leu Ala Gly Gly Gln Leu Ser Leu Lys Ser |      |
| 440 445 450   |      |
| gga gtc act cta gtt gct aag tcc ttt tcg caa tct ccg ggc tct acc | 1507 |
| Gly Val Thr Leu Val Ala Lys Ser Phe Ser Gln Ser Pro Gly Ser Thr |      |
| 455 460 465   |      |
| ctc ctc atg gat gca ggg acc aca tta gaa acc gct gat ggg atc act | 1555 |
| Leu Leu Met Asp Ala Gly Thr Thr Leu Glu Thr Ala Asp Gly Ile Thr |      |
| 470 475 480 485   |      |
| atc aat aat ctt gtt ctc aat gta gat tcc tta aaa gag acc aag aag | 1603 |
| Ile Asn Asn Leu Val Leu Asn Val Asp Ser Leu Lys Glu Thr Lys Lys |      |
| 490 495 500   |      |
| ggc acg cta aaa gca aca caa gca agt cag aca gtc act tta tct gga | 1651 |
| Gly Thr Leu Lys Ala Thr Gln Ala Ser Gln Thr Val Thr Leu Ser Gly |      |
| 505 510 515   |      |
| tcg ctc tct ctt gta gat cct tct gga aat gtc tac gaa gat gtc tct | 1699 |
| Ser Leu Ser Leu Val Asp Pro Ser Gly Asn Val Tyr Glu Asp Val Ser |      |
| 520 525 530   |      |
| tgg aat aac cct caa gtc ttt tct tgt ctc act ctt act gct gac gac | 1747 |
| Trp Asn Asn Pro Gln Val Phe Ser Cys Leu Thr Leu Thr Ala Asp Asp |      |
| 535 540 545   |      |
| ccc gcg aat att cac atc aca gac tta gct gct gat ccc cta gaa aaa | 1795 |
| Pro Ala Asn Ile His Ile Thr Asp Leu Ala Ala Asp Pro Leu Glu Lys |      |
| 550 555 560 565   |      |
| aat cct atc cat tgg gga tac caa ggg aat tgg gca tta tct tgg caa | 1843 |
| Asn Pro Ile His Trp Gly Tyr Gln Gly Asn Trp Ala Leu Ser Trp Gln |      |
| 570 575 580   |      |
| gag gat act gcg act aaa tcc aaa gca gcg act ctt acc tgg aca aaa | 1891 |
| Glu Asp Thr Ala Thr Lys Ser Lys Ala Ala Thr Leu Thr Trp Thr Lys |      |
| 585 590 595   |      |
| aca gga tac aat ccg aat cct gag cgt cgt gga acc tta gtt gct aac | 1939 |
| Thr Gly Tyr Asn Pro Asn Pro Glu Arg Arg Gly Thr Leu Val Ala Asn |      |
| 600 605 610   |      |
| acg cta tgg gga tcc ttt gtt gat gtg cgc tcc ata caa cag ctt gta | 1987 |
| Thr Leu Trp Gly Ser Phe Val Asp Val Arg Ser Ile Gln Gln Leu Val |      |
| 615 620 625   |      |
| gcc act aaa gta cgc caa tct caa gaa act cgc ggc atc tgg tgt gaa | 2035 |
| Ala Thr Lys Val Arg Gln Ser Gln Glu Thr Arg Gly Ile Trp Cys Glu |      |
| 630 635 640 645   |      |

37/77

Fig. 5 (con't)

|   |      |
|---|------|
| ggg atc tcg aac ttc ttc cat aaa gat agc acg aag ata aat aaa ggt | 2083 |
| Gly Ile Ser Asn Phe Phe His Lys Asp Ser Thr Lys Ile Asn Lys Gly |      |
| 650 655 660   |      |
| ttt cgc cac ata agt gca ggt tat gtt gta gga gcg act aca aca tta | 2131 |
| Phe Arg His Ile Ser Ala Gly Tyr Val Val Gly Ala Thr Thr Thr Leu |      |
| 665 670 675   |      |
| gct tct gat aat ctt atc act gca gcc ttc tgc caa tta ttc ggg aaa | 2179 |
| Ala Ser Asp Asn Leu Ile Thr Ala Ala Phe Cys Gln Leu Phe Gly Lys |      |
| 680 685 690   |      |
| gat aga gat cac ttt ata aat aaa aat aga gct tct gcc tat gca gct | 2227 |
| Asp Arg Asp His Phe Ile Asn Lys Asn Arg Ala Ser Ala Tyr Ala Ala |      |
| 695 700 705   |      |
| tct ctc cat ctc cag cat cta gcg acc ttg tct tct cca agc ttg tta | 2275 |
| Ser Leu His Leu Gln His Leu Ala Thr Leu Ser Ser Pro Ser Leu Leu |      |
| 710 715 720 725   |      |
| cgc tac ctt cct gga tct gaa agt gag cag cct gtc ctc ttt gat gct | 2323 |
| Arg Tyr Leu Pro Gly Ser Glu Ser Glu Gln Pro Val Leu Phe Asp Ala |      |
| 730 735 740   |      |
| cag atc agc tat atc tat agt aaa aat act atg aaa acc tat tac acc | 2371 |
| Gln Ile Ser Tyr Ile Tyr Ser Lys Asn Thr Met Lys Thr Tyr Tyr Thr |      |
| 745 750 755   |      |
| caa gca cca aag gga gag agc tcg tgg tat aat gac ggt tgc gct ctg | 2419 |
| Gln Ala Pro Lys Gly Glu Ser Ser Trp Tyr Asn Asp Gly Cys Ala Leu |      |
| 760 765 770   |      |
| gaa ctt gcg agc tcc cta cca cac act gct tta agc cat gag ggt ctc | 2467 |
| Glu Leu Ala Ser Ser Leu Pro His Thr Ala Leu Ser His Glu Gly Leu |      |
| 775 780 785   |      |
| ttc cac gcg tat ttt cct ttc atc aaa gta gaa gct tcg tac ata cac | 2515 |
| Phe His Ala Tyr Phe Pro Phe Ile Lys Val Glu Ala Ser Tyr Ile His |      |
| 790 795 800 805   |      |
| caa gat agc ttc aaa gaa cgt aat act acc ttg gta cga tct ttc gat | 2563 |
| Gln Asp Ser Phe Lys Glu Arg Asn Thr Thr Leu Val Arg Ser Phe Asp |      |
| 810 815 820   |      |
| agc ggt gat tta att aac gtc tct gtg cct att gga att acc ttc gag | 2611 |
| Ser Gly Asp Leu Ile Asn Val Ser Val Pro Ile Gly Ile Thr Phe Glu |      |
| 825 830 835   |      |
| aga ttc tcg aga aac gag cgt gcg tct tac gaa gct act gtc atc tac | 2659 |
| Arg Phe Ser Arg Asn Glu Arg Ala Ser Tyr Glu Ala Thr Val Ile Tyr |      |
| 840 845 850   |      |
| gtt gcc gat gtc tat cgt aag aat cct gac tgc acg aca gct ctc cta | 2707 |
| Val Ala Asp Val Tyr Arg Lys Asn Pro Asp Cys Thr Thr Ala Leu Leu |      |
| 855 860 865   |      |

38/77

Fig. 5 (con't)

```

atc aac aat acc tcg tgg aaa act aca gga acg aat ctc tca aga caa 2755
Ile Asn Asn Thr Ser Trp Lys Thr Thr Gly Thr Asn Leu Ser Arg Gln
870                      875                      880                      885

gct ggt atc gga aga gca ggg atc ttt tat gcc ttc tct cca aat ctt 2803
Ala Gly Ile Gly Arg Ala Gly Ile Phe Tyr Ala Phe Ser Pro Asn Leu
                        890                      895                      900

gag gtc aca agt aac cta tct atg gaa att cgt gga tct tca cgc agc 2851
Glu Val Thr Ser Asn Leu Ser Met Glu Ile Arg Gly Ser Ser Arg Ser
                        905                      910                      915

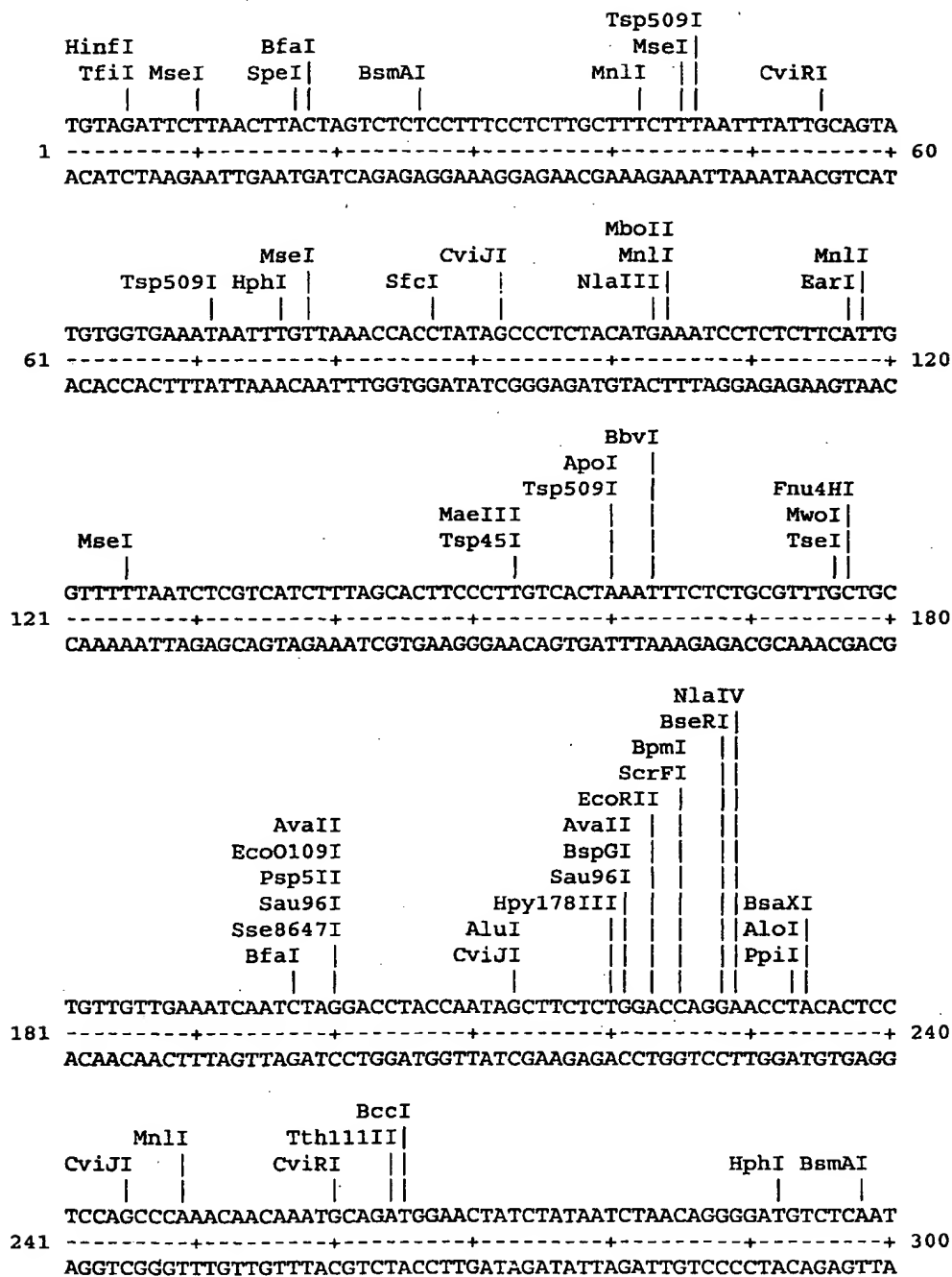
tac aat gca gat ctt gga ggt aag ttc cag ttc taaaagcggt cctgatccct 2904
Tyr Asn Ala Asp Leu Gly Gly Lys Phe Gln Phe
      920                      925

tagaaattct aagagatcct gagtgtatct agggacttct caaaga 2950

```

39/77

Figure 6: Restriction enzyme analysis of CPN100638



**Fig. 6 (con't)**



41/77

Fig. 6 (con't)

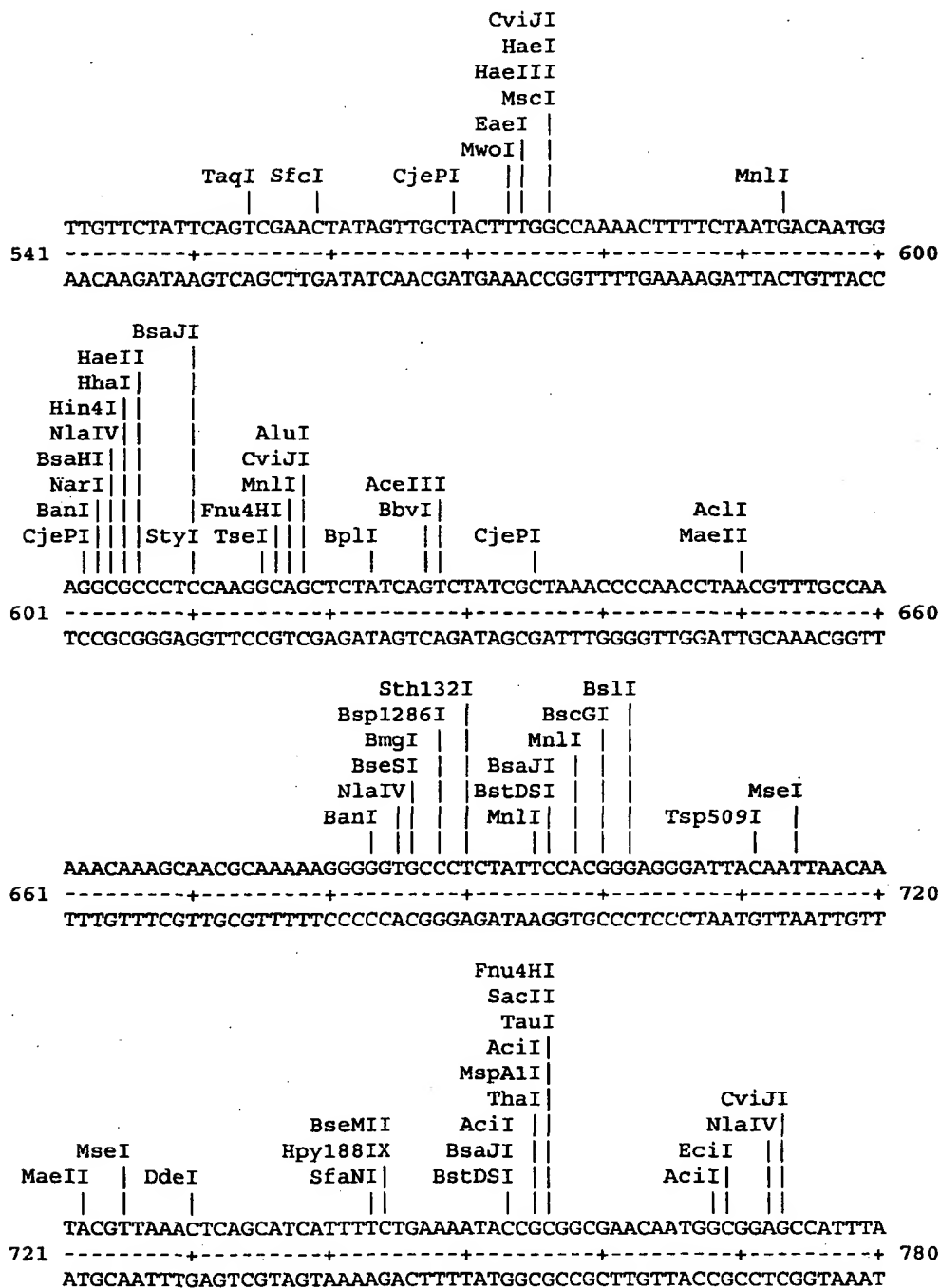




Fig. 6 (con't)



Fig. 6 (con't)

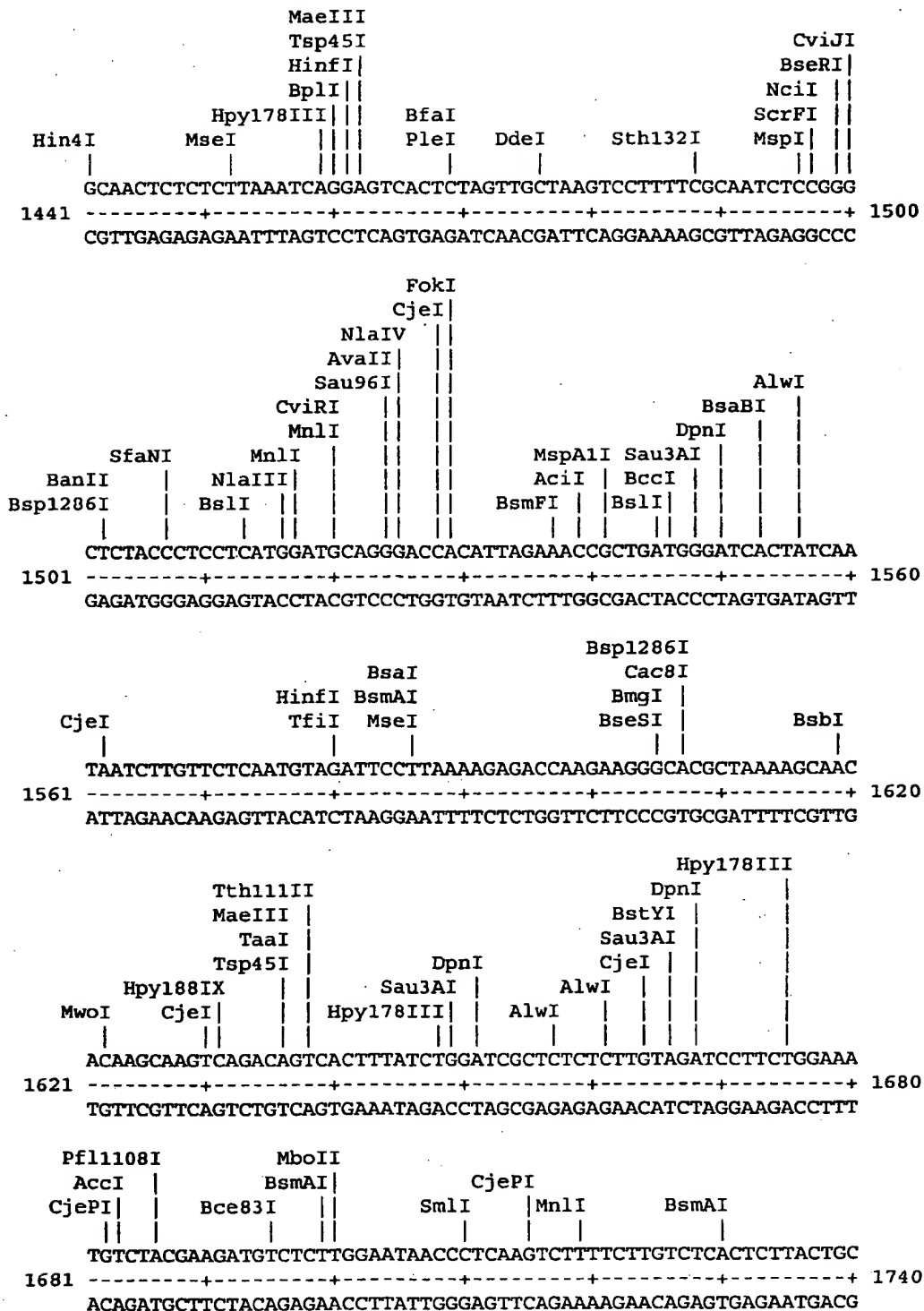


Fig. 6 (con't)



45/77

Fig. 6 (con't)

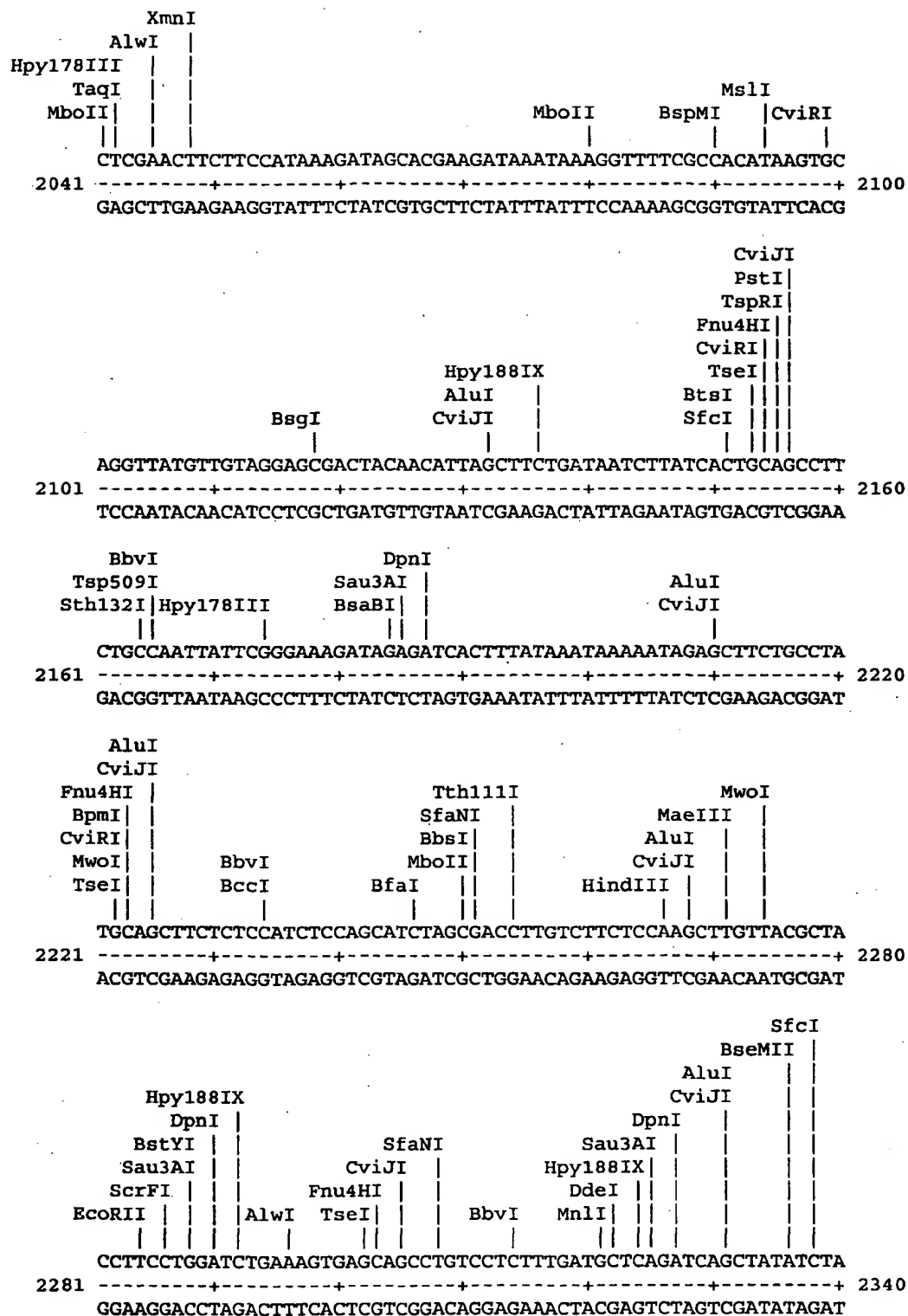


**Fig. 6 (con't)**



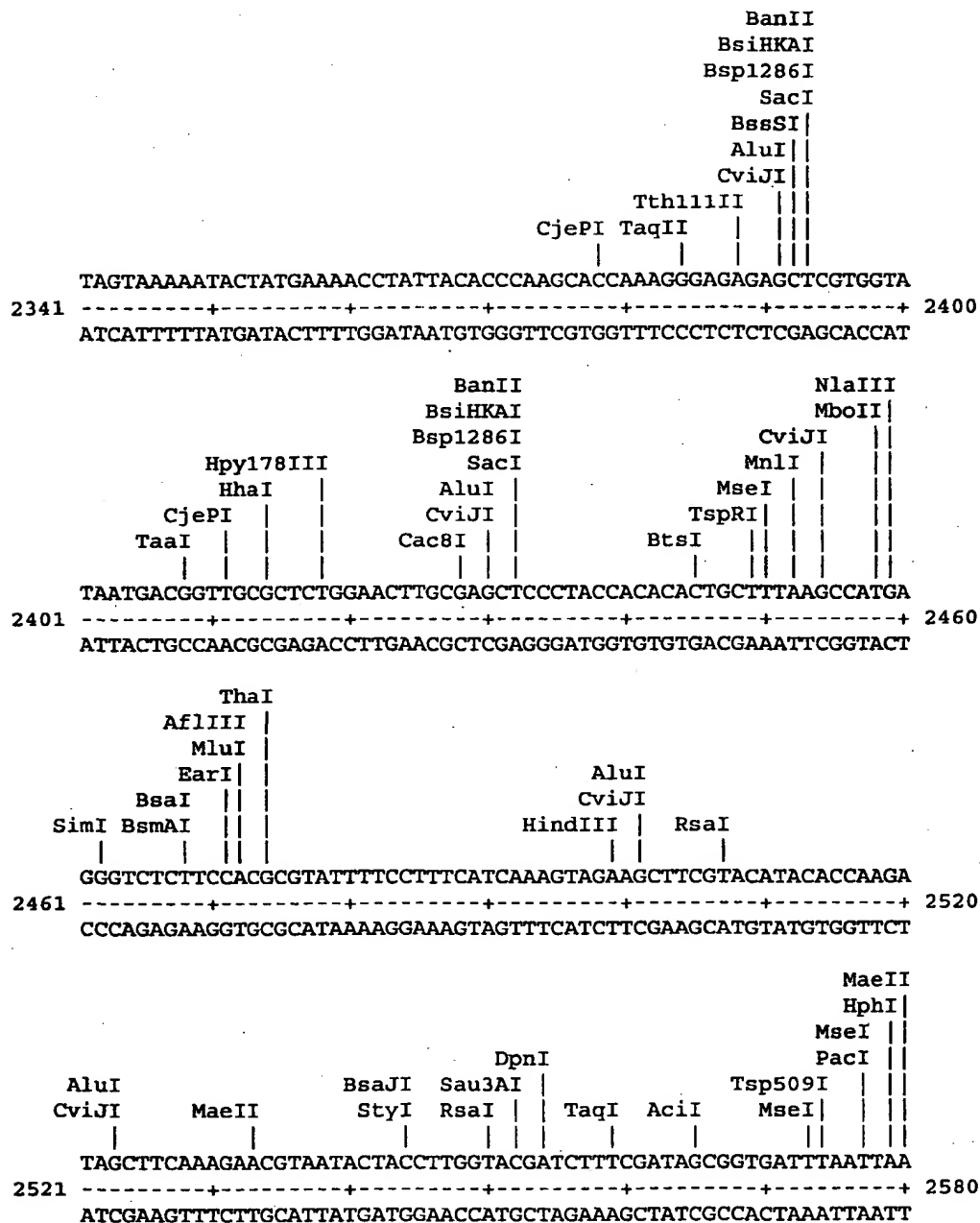
47/77

Fig. 6 (con't)



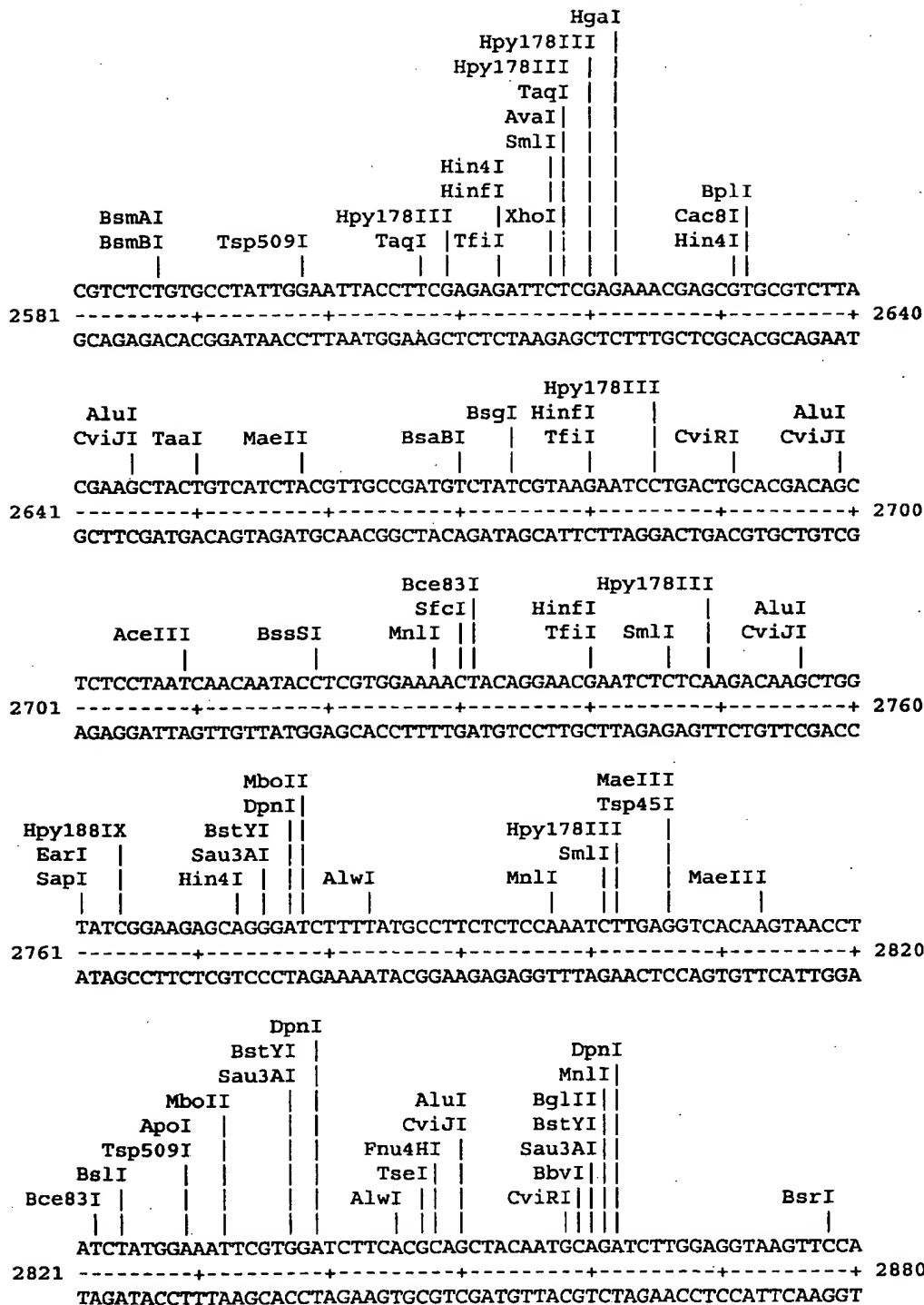
48/77

Fig. 6 (con't)



49/77

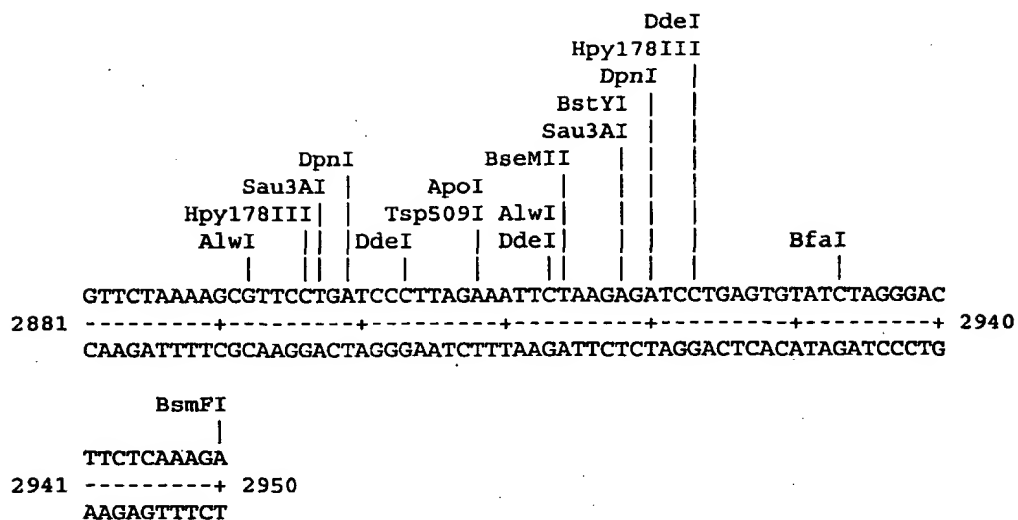
Fig. 6 (con't)





50/77

Fig. 6 (con't)



51/77

Figure 7: Nucleotide and amino acid sequence of CPN100639

```

gtacgaagtt cttcacgaaa ttataataca aacctaggct ctaagttttg tttctagatt 60
atcgaaaacg tgtaattaa ttgaacccaa gcacctttct atg aaa ata ccc ttg 115
                                     Met Lys Ile Pro Leu
                                     1           5
cac aaa ctc ctg atc tct tcg act ctt gtc act ccc att cta ttg agc 163
His Lys Leu Leu Ile Ser Ser Thr Leu Val Thr Pro Ile Leu Leu Ser
                                     10           15           20
att gca act tac gga gca gat gct tct tta tcc cct aca gat agc ttt 211
Ile Ala Thr Tyr Gly Ala Asp Ala Ser Leu Ser Pro Thr Asp Ser Phe
                                     25           30           35
gat gga gcg ggc ggc tct aca ttt act cca aaa tct aca gca gat gcc 259
Asp Gly Ala Gly Gly Ser Thr Phe Thr Pro Lys Ser Thr Ala Asp Ala
                                     40           45           50
aat gga acg aac tat gtc tta tca gga aat gtc tat ata aac gat gct 307
Asn Gly Thr Asn Tyr Val Leu Ser Gly Asn Val Tyr Ile Asn Asp Ala
                                     55           60           65
ggg aaa ggc aca gca tta aca ggc tgc tgc ttt aca gaa act acg ggt 355
Gly Lys Gly Thr Ala Leu Thr Gly Cys Cys Phe Thr Glu Thr Thr Gly
                                     70           75           80           85
- gat ctg aca ttt act gga aag gga tac tca ttt tca ttc aac acg gta 403
Asp Leu Thr Phe Thr Gly Lys Gly Tyr Ser Phe Ser Phe Asn Thr Val
                                     90           95           100
gat gcg ggt tcg aat gca gga gct gcg gca agc aca act gct gat aaa 451
Asp Ala Gly Ser Asn Ala Gly Ala Ala Ala Ser Thr Thr Ala Asp Lys
                                     105           110           115
gcc cta atc ttc aca gga ttt tct aac ctt tcc ttc att gca gct cct 499
Ala Leu Ile Phe Thr Gly Phe Ser Asn Leu Ser Phe Ile Ala Ala Pro
                                     120           125           130
gga act aca gtt gct tca gga aaa agt act tta agt tct gca gga gcc 547
Gly Thr Thr Val Ala Ser Gly Lys Ser Thr Leu Ser Ser Ala Gly Ala
                                     135           140           145
tta aat ctt acc gat aat gga acg att ctc ttt agc caa aac gtc tcc 595
Leu Asn Leu Thr Asp Asn Gly Thr Ile Leu Phe Ser Gln Asn Val Ser
                                     150           155           160           165
aat gaa gct aat aac aat ggc gga gcg atc acc aca aaa act ctt tct 643
Asn Glu Ala Asn Asn Asn Gly Gly Ala Ile Thr Thr Lys Thr Leu Ser
                                     170           175           180
att tct ggg aat acc tct tct ata acc ttc act agt aat agc gca aaa 691
Ile Ser Gly Asn Thr Ser Ser Ile Thr Phe Thr Ser Asn Ser Ala Lys
                                     185           190           195

```

52/77

Fig. 7 (con't)

|   |      |
|---|------|
| aaa tta ggt gga gcg atc tat agc tct gcg gct gca agt att tca gga | 739  |
| Lys Leu Gly Gly Ala Ile Tyr Ser Ser Ala Ala Ala Ser Ile Ser Gly |      |
| 200 205 210   |      |
| aac acc ggc cag tta gtc ttt atg aat aat aaa gga gaa act ggg ggt | 787  |
| Asn Thr Gly Gln Leu Val Phe Met Asn Asn Lys Gly Glu Thr Gly Gly |      |
| 215 220 225   |      |
| ggg gct ctg ggc ttt gaa gcc agc tcc tcg att act caa aat agc tcc | 835  |
| Gly Ala Leu Gly Phe Glu Ala Ser Ser Ser Ile Thr Gln Asn Ser Ser |      |
| 230 235 240 245   |      |
| ctt ttc ttc tct gga aac act gca aca gat gct gca ggc aag ggc ggg | 883  |
| Leu Phe Phe Ser Gly Asn Thr Ala Thr Asp Ala Ala Gly Lys Gly Gly |      |
| 250 255 260   |      |
| gcc att tat tgt gaa aaa aca gga gag act cct act ctt act atc tct | 931  |
| Ala Ile Tyr Cys Glu Lys Thr Gly Glu Thr Pro Thr Leu Thr Ile Ser |      |
| 265 270 275   |      |
| gga aat aaa agt ctg acc ttc gcc gag aac tct tca gta act caa ggc | 979  |
| Gly Asn Lys Ser Leu Thr Phe Ala Glu Asn Ser Ser Val Thr Gln Gly |      |
| 280 285 290   |      |
| gga gca atc tgt gcc cat ggt cta gat ctt tcc gct gct ggc cct acc | 1027 |
| Gly Ala Ile Cys Ala His Gly Leu Asp Leu Ser Ala Ala Gly Pro Thr |      |
| 295 300 305   |      |
| cta ttt tca aat aat aga tgc ggg aac aca gct gca ggc aag ggc ggc | 1075 |
| Leu Phe Ser Asn Asn Arg Cys Gly Asn Thr Ala Ala Gly Lys Gly Gly |      |
| 310 315 320 325   |      |
| gct att gca att gcc gac tct gga tct tta agt ctc tct gca aat caa | 1123 |
| Ala Ile Ala Ile Ala Asp Ser Gly Ser Leu Ser Leu Ser Ala Asn Gln |      |
| 330 335 340   |      |
| gga gac atc acg ttc ctt ggc aac act cta acc tca acc tcc gcg cca | 1171 |
| Gly Asp Ile Thr Phe Leu Gly Asn Thr Leu Thr Ser Thr Ser Ala Pro |      |
| 345 350 355   |      |
| aca tcg aca cgg aat gct atc tac ctg gga tcg tca gca aaa att acg | 1219 |
| Thr Ser Thr Arg Asn Ala Ile Tyr Leu Gly Ser Ser Ala Lys Ile Thr |      |
| 360 365 370   |      |
| aac tta agg gca gcc caa ggc caa tct atc tat ttc tat gat ccg att | 1267 |
| Asn Leu Arg Ala Ala Gln Gly Gln Ser Ile Tyr Phe Tyr Asp Pro Ile |      |
| 375 380 385   |      |
| gca tct aac acc aca gga gct tca gac gtt ctg acc atc aac caa ccg | 1315 |
| Ala Ser Asn Thr Thr Gly Ala Ser Asp Val Leu Thr Ile Asn Gln Pro |      |
| 390 395 400 405   |      |
| gat agc aac tcg cct tta gat tat tca gga acg att gta ttt tct ggg | 1363 |
| Asp Ser Asn Ser Pro Leu Asp Tyr Ser Gly Thr Ile Val Phe Ser Gly |      |
| 410 415 420   |      |

53/77

Fig. 7 (con't)

|   |      |
|---|------|
| gaa aag ctc tct gca gat gaa gcg aaa gct gct gat aac ttc aca tct | 1411 |
| Glu Lys Leu Ser Ala Asp Glu Ala Lys Ala Ala Asp Asn Phe Thr Ser |      |
| 425 430 435   |      |
| ata tta aag caa cca ttg gct cta gcc tct gga acc tta gca ctc aaa | 1459 |
| Ile Leu Lys Gln Pro Leu Ala Leu Ala Ser Gly Thr Leu Ala Leu Lys |      |
| 440 445 450   |      |
| gga aat gtc gag tta gat gtc aat ggt ttc aca cag act gaa ggc tct | 1507 |
| Gly Asn Val Glu Leu Asp Val Asn Gly Phe Thr Gln Thr Glu Gly Ser |      |
| 455 460 465   |      |
| aca ctc ctc atg caa cca gga aca aag ctc aaa gca gat act gaa gct | 1555 |
| Thr Leu Leu Met Gln Pro Gly Thr Lys Leu Lys Ala Asp Thr Glu Ala |      |
| 470 475 480 485   |      |
| atc agt ctt acc aaa ctt gtc gtt gat ctt tct gcc tta gag gga aat | 1603 |
| Ile Ser Leu Thr Lys Leu Val Val Asp Leu Ser Ala Leu Glu Gly Asn |      |
| 490 495 500   |      |
| aag agt gtg tcc att gaa aca gca gga gcc aac aaa act ata act cta | 1651 |
| Lys Ser Val Ser Ile Glu Thr Ala Gly Ala Asn Lys Thr Ile Thr Leu |      |
| 505 510 515   |      |
| acc tct cct ctt gtt ttc caa gat agt agc ggc aat ttt tat gaa agc | 1699 |
| Thr Ser Pro Leu Val Phe Gln Asp Ser Ser Gly Asn Phe Tyr Glu Ser |      |
| 520 525 530   |      |
| cat acg ata aac caa gcc ttc acg cag cct ttg gtg gta ttc act gct | 1747 |
| His Thr Ile Asn Gln Ala Phe Thr Gln Pro Leu Val Val Phe Thr Ala |      |
| 535 540 545   |      |
| gct act gct gct agc gat att tat atc gat gcg ctt ctc act tct cca | 1795 |
| Ala Thr Ala Ala Ser Asp Ile Tyr Ile Asp Ala Leu Leu Thr Ser Pro |      |
| 550 555 560 565   |      |
| gta caa act cca gaa cct cat tac ggg tat cag gga cat tgg gaa gcc | 1843 |
| Val Gln Thr Pro Glu Pro His Tyr Gly Tyr Gln Gly His Trp Glu Ala |      |
| 570 575 580   |      |
| act tgg gca gac aca tca act gca aaa tca gga act atg act tgg gta | 1891 |
| Thr Trp Ala Asp Thr Ser Thr Ala Lys Ser Gly Thr Met Thr Trp Val |      |
| 585 590 595   |      |
| act acg ggc tac aac cct aat cct gag cgt aga gct tcc gta gtt ccc | 1939 |
| Thr Thr Gly Tyr Asn Pro Asn Pro Glu Arg Arg Ala Ser Val Val Pro |      |
| 600 605 610   |      |
| gat tca tta tgg gca tcc ttt act gac att cgc act cta cag cag atc | 1987 |
| Asp Ser Leu Trp Ala Ser Phe Thr Asp Ile Arg Thr Leu Gln Gln Ile |      |
| 615 620 625   |      |
| atg aca tct caa gcg aat agt atc tat cag caa cga gga ctc tgg gca | 2035 |
| Met Thr Ser Gln Ala Asn Ser Ile Tyr Gln Gln Arg Gly Leu Trp Ala |      |
| 630 635 640 645   |      |

54/77

Fig. 7 (con't)

|   |      |
|---|------|
| tca gga act gcg aat ttc ttc cat aag gat aaa tca gga act aac caa | 2083 |
| Ser Gly Thr Ala Asn Phe Phe His Lys Asp Lys Ser Gly Thr Asn Gln |      |
| 650 655 660   |      |
| gca ttc cga cat aaa agc tac ggc tat att gtt gga gga agt gct gaa | 2131 |
| Ala Phe Arg His Lys Ser Tyr Gly Tyr Ile Val Gly Gly Ser Ala Glu |      |
| 665 670 675   |      |
| gat ttt tct gaa aat atc ttc agt gta gct ttc tgc cag ctc ttc ggt | 2179 |
| Asp Phe Ser Glu Asn Ile Phe Ser Val Ala Phe Cys Gln Leu Phe Gly |      |
| 680 685 690   |      |
| aaa gat aaa gac ctg ttt ata gtt gaa aat acc tct cat aac tat tta | 2227 |
| Lys Asp Lys Asp Leu Phe Ile Val Glu Asn Thr Ser His Asn Tyr Leu |      |
| 695 700 705   |      |
| gcg tcg cta tac ctg caa cat cga gca ttc cta gga gga ctt ccc atg | 2275 |
| Ala Ser Leu Tyr Leu Gln His Arg Ala Phe Leu Gly Gly Leu Pro Met |      |
| 710 715 720 725   |      |
| ccc tca ttt gga agt atc acc gac atg ctg aaa gat att cct ctc att | 2323 |
| Pro Ser Phe Gly Ser Ile Thr Asp Met Leu Lys Asp Ile Pro Leu Ile |      |
| 730 735 740   |      |
| ttg aat gcc cag cta agc tac agc tac act aaa aat gat atg gat act | 2371 |
| Leu Asn Ala Gln Leu Ser Tyr Ser Tyr Thr Lys Asn Asp Met Asp Thr |      |
| 745 750 755   |      |
| cgc tat act tcc tat cct gaa gct caa ggc tct tgg acc aat aac tct | 2419 |
| Arg Tyr Thr Ser Tyr Pro Glu Ala Gln Gly Ser Trp Thr Asn Asn Ser |      |
| 760 765 770   |      |
| ggg gct cta gag ctc gga gga tct ctg gct cta tat ctc cct aaa gaa | 2467 |
| Gly Ala Leu Glu Leu Gly Gly Ser Leu Ala Leu Tyr Leu Pro Lys Glu |      |
| 775 780 785   |      |
| gca ccg ttc ttc cag gga tat ttc ccc ttc tta aag ttc cag gca gtc | 2515 |
| Ala Pro Phe Phe Gln Gly Tyr Phe Pro Phe Leu Lys Phe Gln Ala Val |      |
| 790 795 800 805   |      |
| tac agc cgc caa caa aac ttt aaa gag agt ggc gct gaa gcc cgt gct | 2563 |
| Tyr Ser Arg Gln Gln Asn Phe Lys Glu Ser Gly Ala Glu Ala Arg Ala |      |
| 810 815 820   |      |
| ttt gat gat gga gac cta gtg aac tgc tct atc cct gtc ggc att cgg | 2611 |
| Phe Asp Asp Gly Asp Leu Val Asn Cys Ser Ile Pro Val Gly Ile Arg |      |
| 825 830 835   |      |
| tta gaa aaa atc tcc gaa gat gaa aaa aat aat ttc gag att tct cta | 2659 |
| Leu Glu Lys Ile Ser Glu Asp Glu Lys Asn Asn Phe Glu Ile Ser Leu |      |
| 840 845 850   |      |
| gcc tac att ggt gat gtg tat cgt aaa aat ccc cgt tcg cgt act tct | 2707 |
| Ala Tyr Ile Gly Asp Val Tyr Arg Lys Asn Pro Arg Ser Arg Thr Ser |      |
| 855 860 865   |      |

55/77

Fig. 7 (con't)

```

cta atg gtc agt gga gcc tct tgg act tcg cta tgt aaa aac ctc gca 2755
Leu Met Val Ser Gly Ala Ser Trp Thr Ser Leu Cys Lys Asn Leu Ala
870                875                880                885

cga caa gcc ttc tta gca agt gct gga agc cat ctg act ctc tcc cct 2803
Arg Gln Ala Phe Leu Ala Ser Ala Gly Ser His Leu Thr Leu Ser Pro
890                895                900

cat gta gaa ctc tct ggg gaa gct gct tat gag ctt cgt ggc tca gca 2851
His Val Glu Leu Ser Gly Glu Ala Ala Tyr Glu Leu Arg Gly Ser Ala
905                910                915

cac atc tac aat gta gat tgt ggg cta aga tac tca ttc tagttcctac 2900
His Ile Tyr Asn Val Asp Cys Gly Leu Arg Tyr Ser Phe
920                925                930

tttcctccct aaacttttag ggaggaattc ttataaaaac cctgtagatt cttaacttac 2960

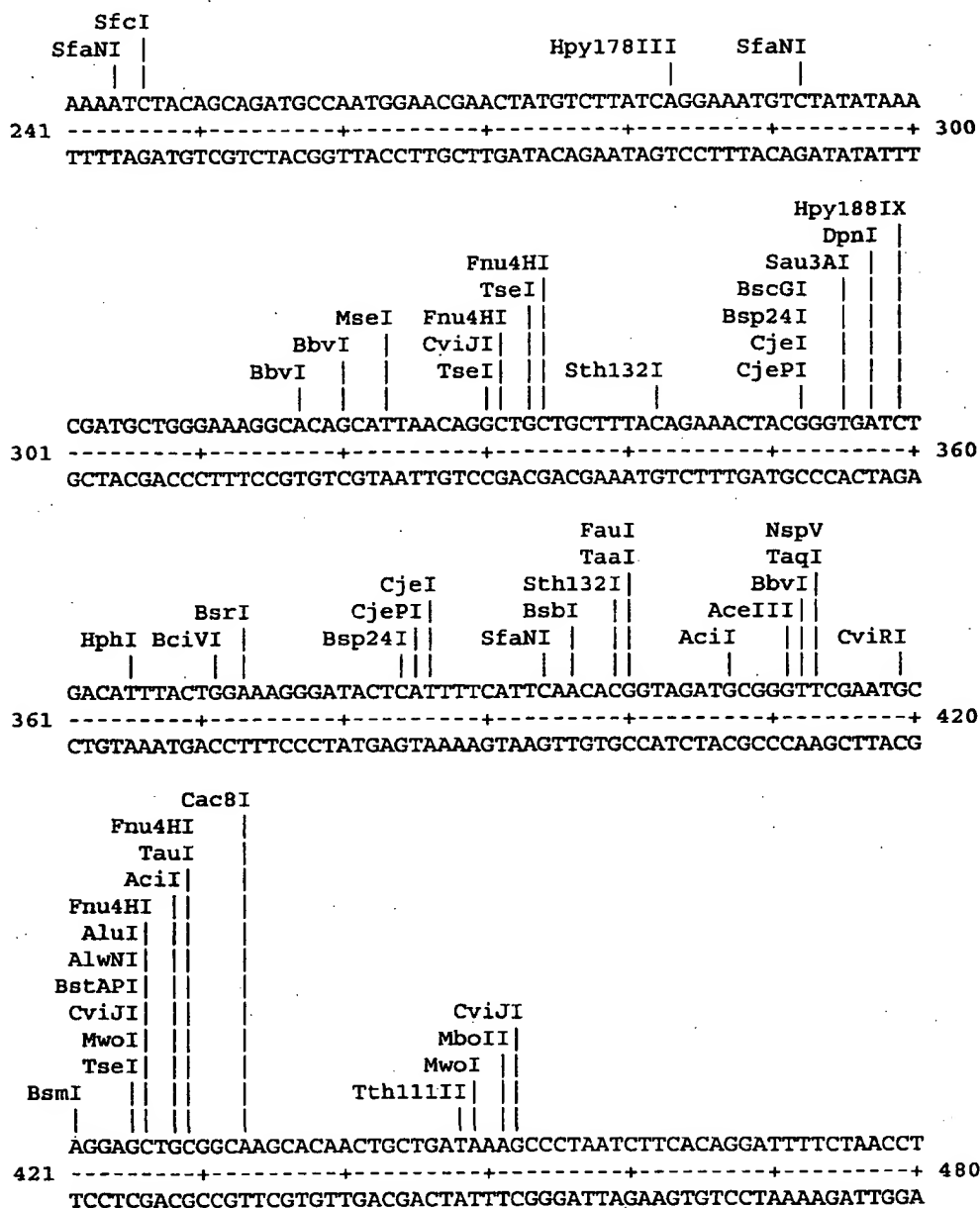
tagtctctcc tttcctcttg ctttctttaa tttattgcag 3000

```



57/77

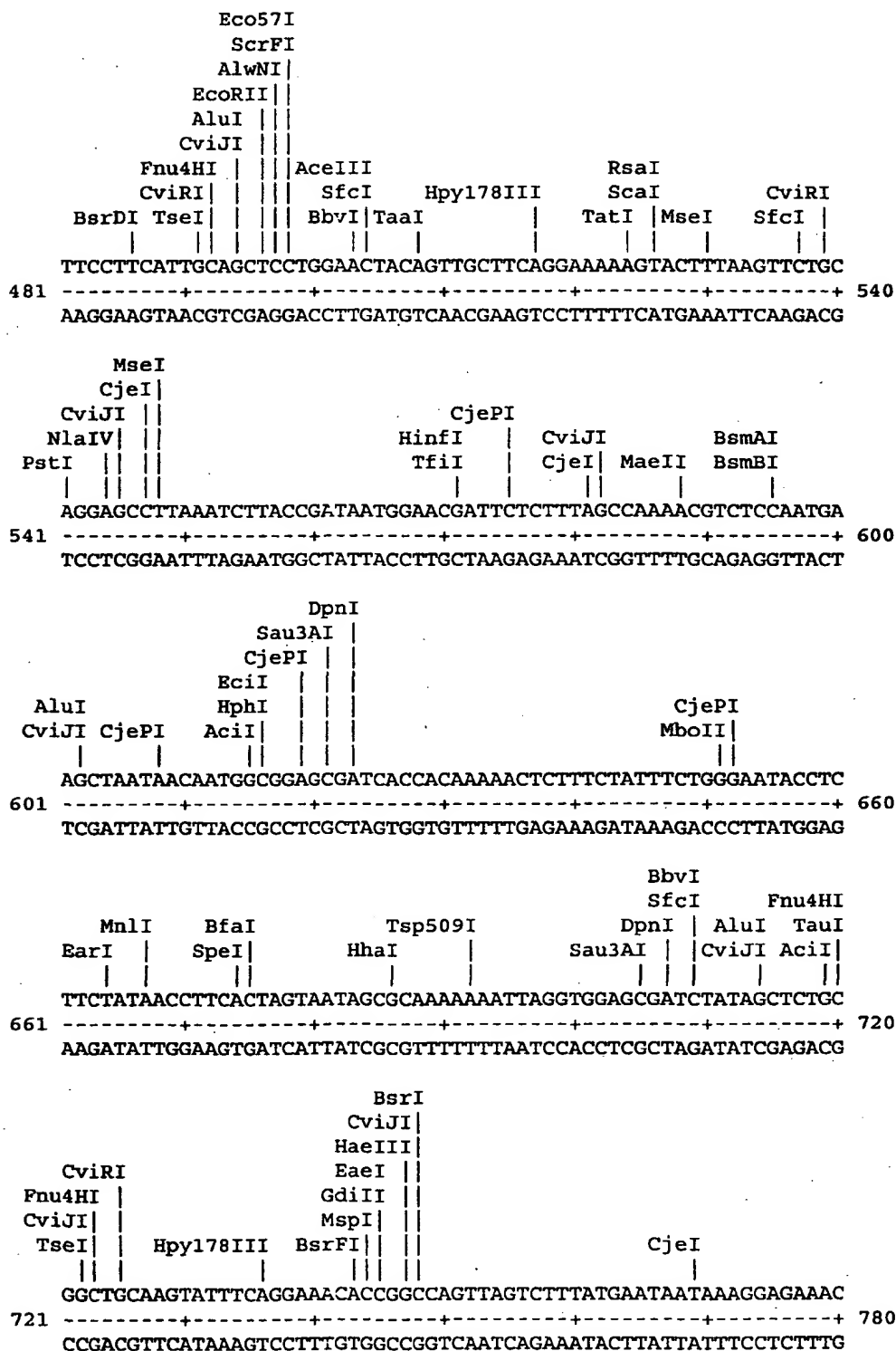
Fig. 8 (con't)





58/77

Fig. 8 (con't)



**Fig. 8 (con't)**

**SUBSTITUTE SHEET (RULE 26)**

Fig. 8 (con't)

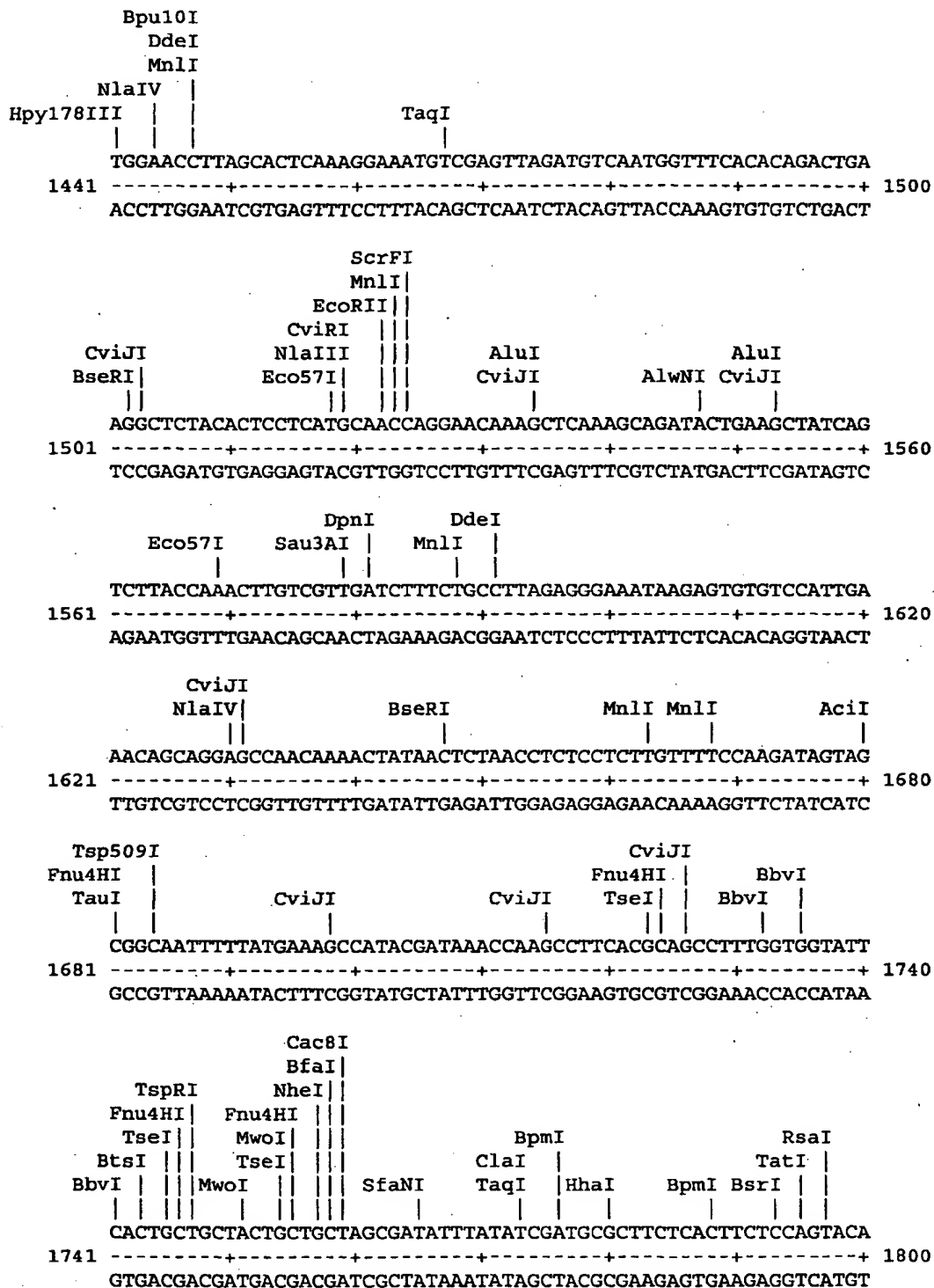


Fig. 8 (con't)



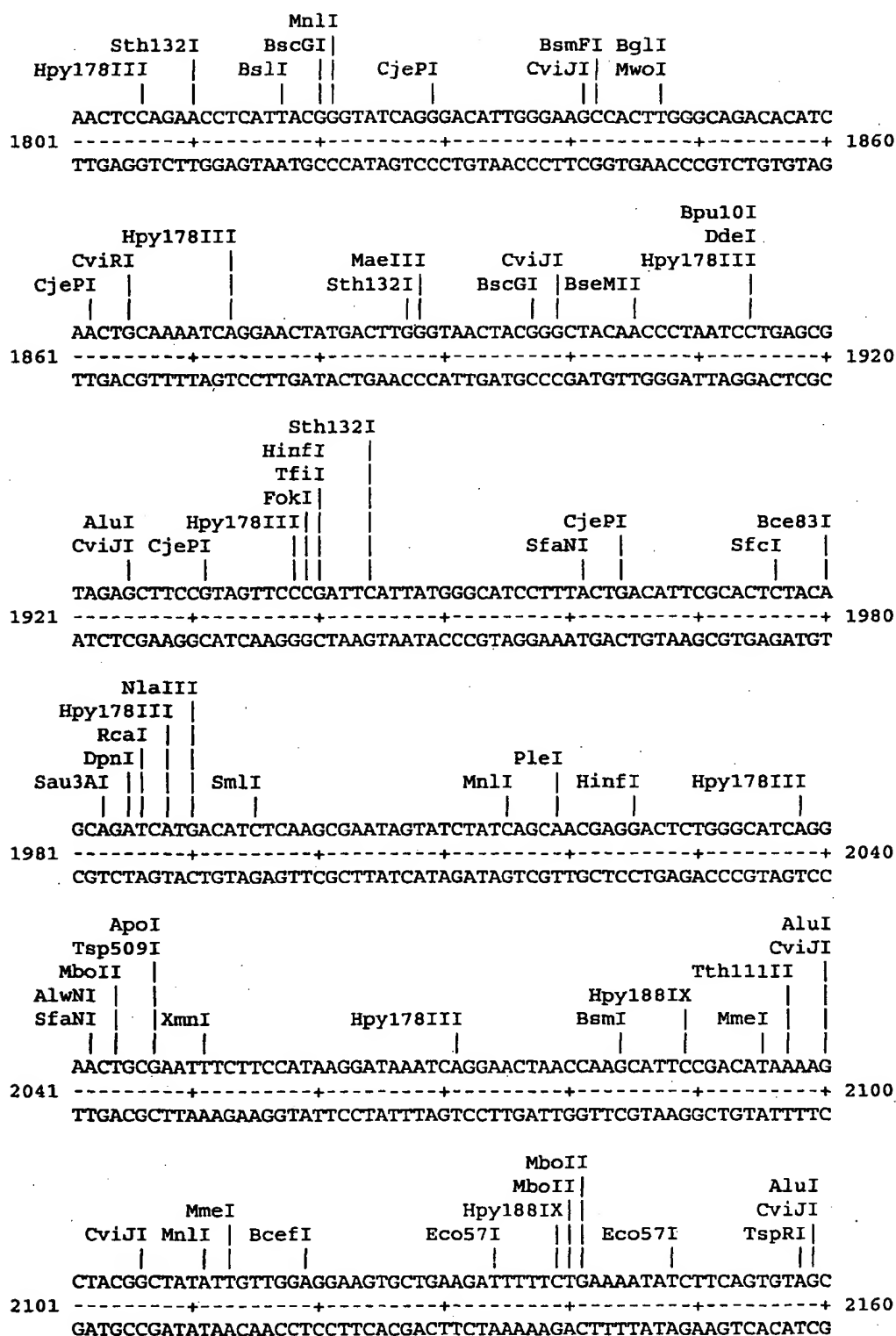
62/77

Fig. 8 (con't)



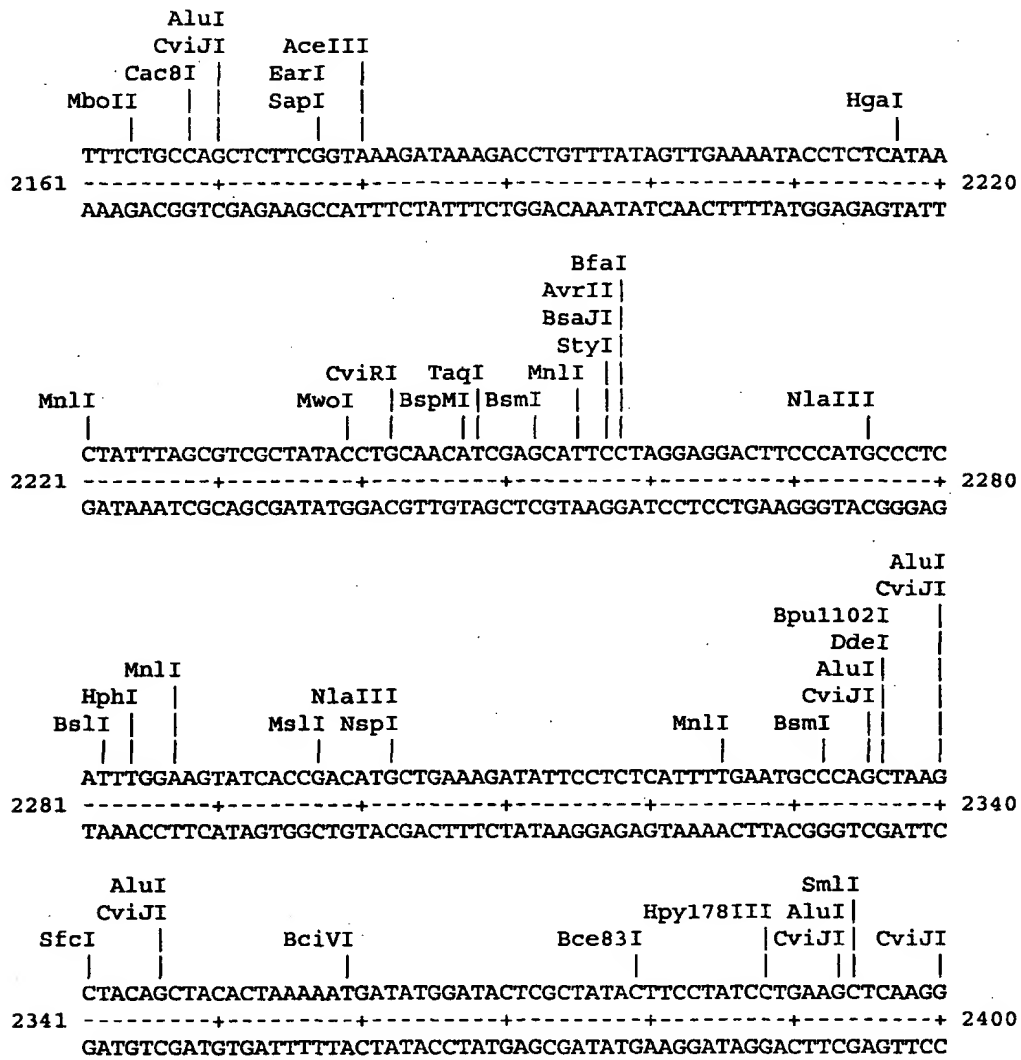
63/77

Fig. 8 (con't)



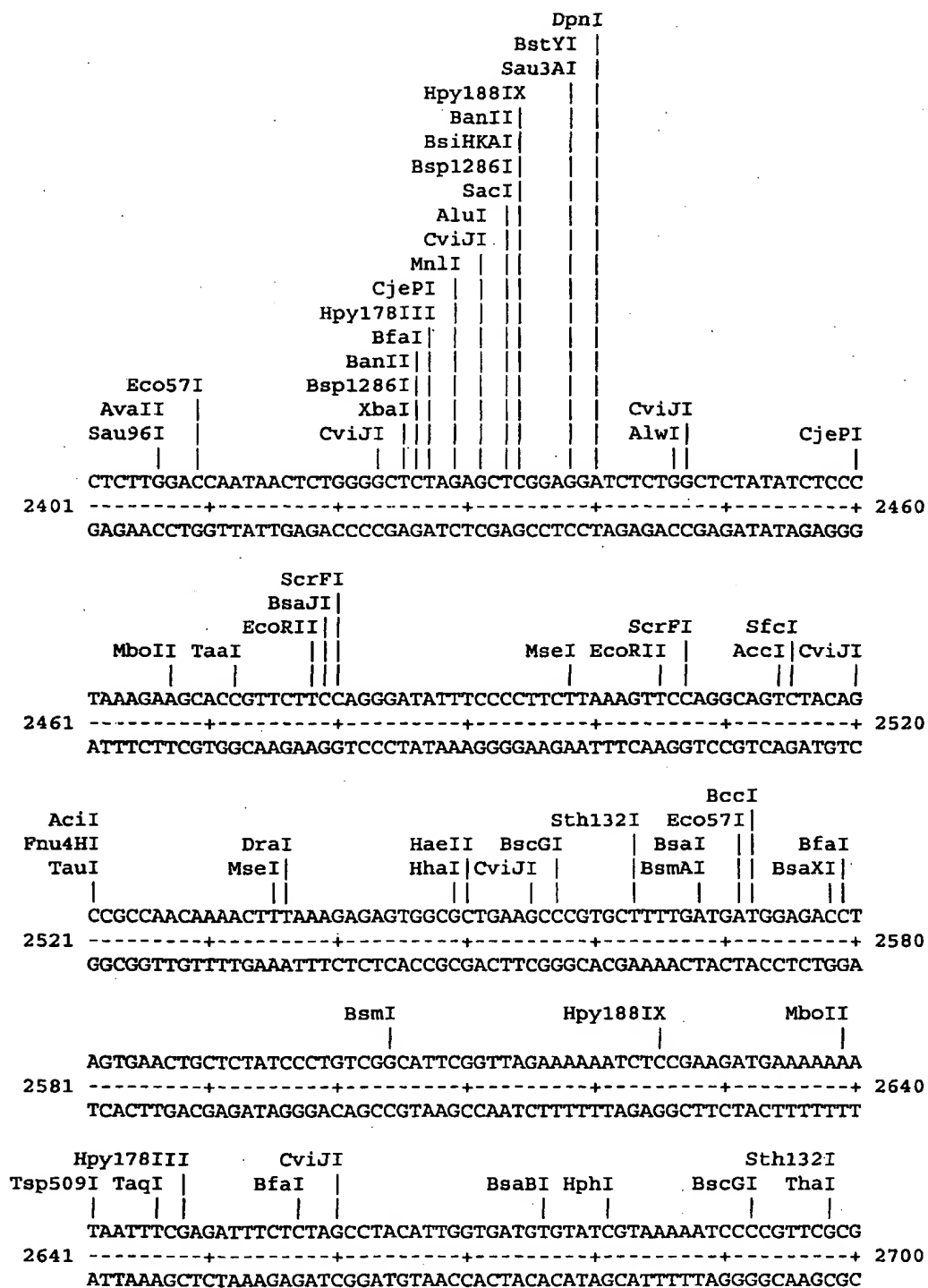
64/77

Fig. 8 (con't)



65/77

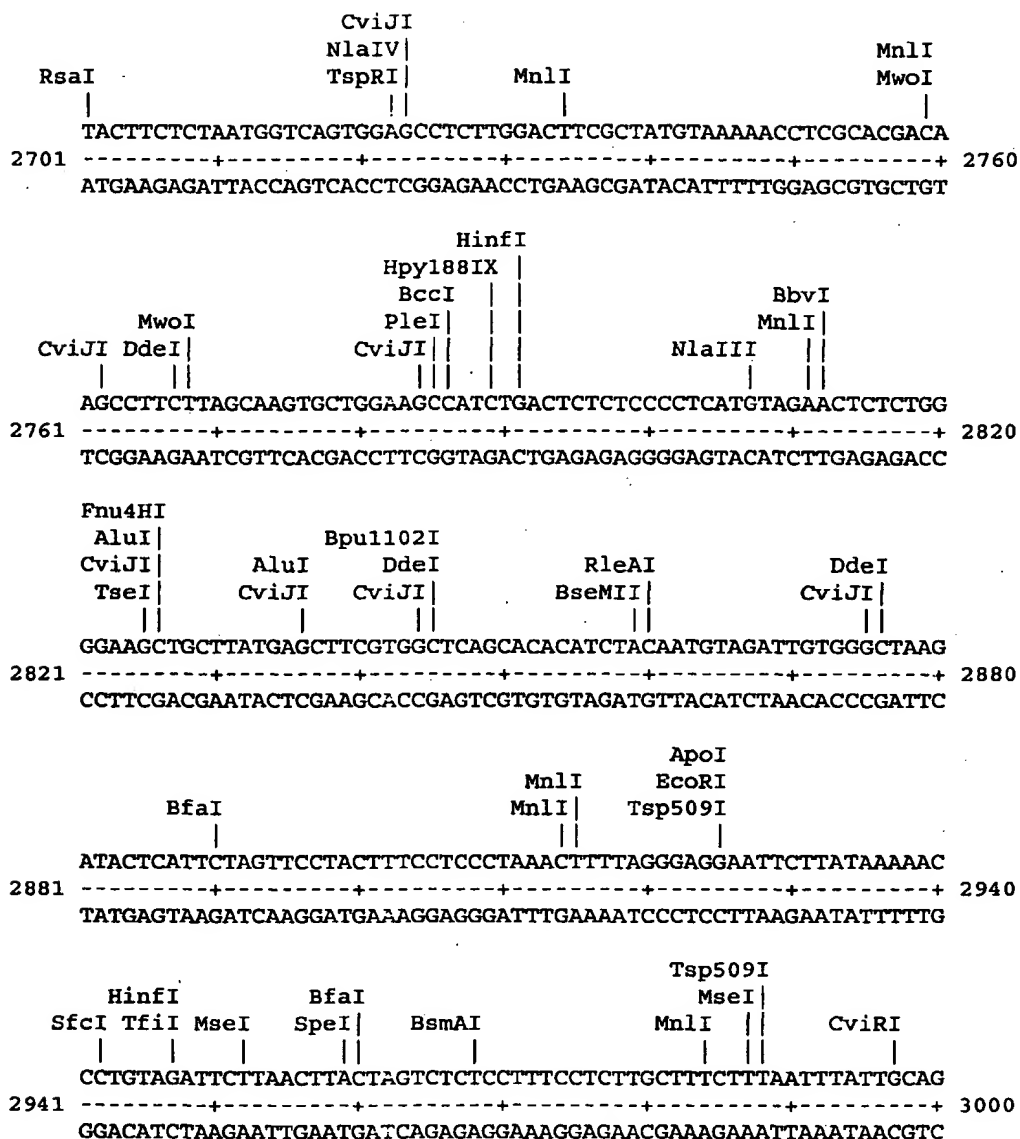
Fig. 8 (con't)





66/77

Fig. 8 (con't)



67/77

Figure 9: Nucleotide and amino acid sequence of CPN100708

```

ataaagtttt ttatatgaac aaaactttga cttcattggc agacacttcg accttaacag 60

accaattttg ttgtcatccc tataaaaatc aggaattttc atg ctc tcc tca cta 115
                                   Met Leu Ser Ser Leu
                                   1       5

atc cgt gat tca ttt ccc ctt ctt att tta ctt ccc aca ttc cta gcg 163
Ile Arg Asp Ser Phe Pro Leu Leu Ile Leu Leu Pro Thr Phe Leu Ala
                                   10       15       20

gca tta gga gcc tcc gta gct ggc ggc gtt atg gga acc tat atc gtt 211
Ala Leu Gly Ala Ser Val Ala Gly Gly Val Met Gly Thr Tyr Ile Val
                                   25       30       35

gta aaa cgt att gtt tca att agt gga agt ata tct cat gca att cta 259
Val Lys Arg Ile Val Ser Ile Ser Gly Ser Ile Ser His Ala Ile Leu
                                   40       45       50

gga gga att ggc ctc acc cta tgg ata caa tat aag ctt cat ctc tct 307
Gly Gly Ile Gly Leu Thr Leu Trp Ile Gln Tyr Lys Leu His Leu Ser
                                   55       60       65

ttt ttc cct atg tat gga gct att gta gga gct att ttt cta gct ctt 355
Phe Phe Pro Met Tyr Gly Ala Ile Val Gly Ala Ile Phe Leu Ala Leu
                                   70       75       80       85

tgc atc ggc aag atc cac ctg aaa tac caa gaa agg gaa gac tct ttg 403
Cys Ile Gly Lys Ile His Leu Lys Tyr Gln Glu Arg Glu Asp Ser Leu
                                   90       95       100

att gcg atg att tgg tct gtg ggc atg gca att gga att ata ttc att 451
Ile Ala Met Ile Trp Ser Val Gly Met Ala Ile Gly Ile Ile Phe Ile
                                   105       110       115

tcc agg ctt ccc acc ttt aat gga gag ctc atc aat ttt cta ttt ggg 499
Ser Arg Leu Pro Thr Phe Asn Gly Glu Leu Ile Asn Phe Leu Phe Gly
                                   120       125       130

aac att ctc tgg gtc acc cct tca gac ctc tat agc tta gga atc ttt 547
Asn Ile Leu Trp Val Thr Pro Ser Asp Leu Tyr Ser Leu Gly Ile Phe
                                   135       140       145

gat ctt ctt gtt tta gga att gtg gtc ctt tgc cac acc cgg ttc ctt 595
Asp Leu Leu Val Leu Gly Ile Val Val Leu Cys His Thr Arg Phe Leu
                                   150       155       160       165

gct ctt tgc ttt gat gag agg tac acg gct tta aac cat tgt tct gta 643
Ala Leu Cys Phe Asp Glu Arg Tyr Thr Ala Leu Asn His Cys Ser Val
                                   170       175       180

cag ctg tgg tat ttc cta ctt ctt gtt ctg aca gca atc acg att gtg 691
Gln Leu Trp Tyr Phe Leu Leu Leu Val Leu Thr Ala Ile Thr Ile Val
                                   185       190       195

```

SUBSTITUTE SHEET (RULE 26)

68/77

Fig. 9 (con't)

|   |      |
|---|------|
| atg ttg att tat gtg atg gga acg att ctg atg ctt agc atg ctc gtc   | 739  |
| Met Leu Ile Tyr Val Met Gly Thr Ile Leu Met Leu Ser Met Leu Val   |      |
| 200 205 210   |      |
| tta cct gtt gct ata gcg tgt aga ttt tcg tac aag atg aca cga att   | 787  |
| Leu Pro Val Ala Ile Ala Cys Arg Phe Ser Tyr Lys Met Thr Arg Ile   |      |
| 215 220 225   |      |
| atg ttc atc tcg gtc ctc ttg aat atc tta tgt tct ttt tct gga att   | 835  |
| Met Phe Ile Ser Val Leu Leu Asn Ile Leu Cys Ser Phe Ser Gly Ile   |      |
| 230 235 240 245   |      |
| tgc atc gcc tac tgt tta gat ttc cca gta ggt cct acg ata tca ttg   | 883  |
| Cys Ile Ala Tyr Cys Leu Asp Phe Pro Val Gly Pro Thr Ile Ser Leu   |      |
| 250 255 260   |      |
| ctg atg ggg tta ggt tat aca gcg agt ctt tgt gtg aag aag cgg tac   | 931  |
| Leu Met Gly Leu Gly Tyr Thr Ala Ser Leu Cys Val Lys Lys Arg Tyr   |      |
| 265 270 275   |      |
| aat ccg tcg acg cct tct cct gta agt cct gaa atc aat aca aat gta   | 979  |
| Asn Pro Ser Thr Pro Ser Pro Val Ser Pro Glu Ile Asn Thr Asn Val   |      |
| 280 285 290   |      |
| tagctagggg agcgcttttg gaagctttgg aggcattctt cctgttcgtc aggaagaaga | 1039 |
| tcatcaattt tatttaaagc taccagcata tctttctttt caaaatctgg ctgatgagag | 1099 |
| t   | 1100 |



Fig. 10 (con't)



71/77

Fig. 10 (con't)

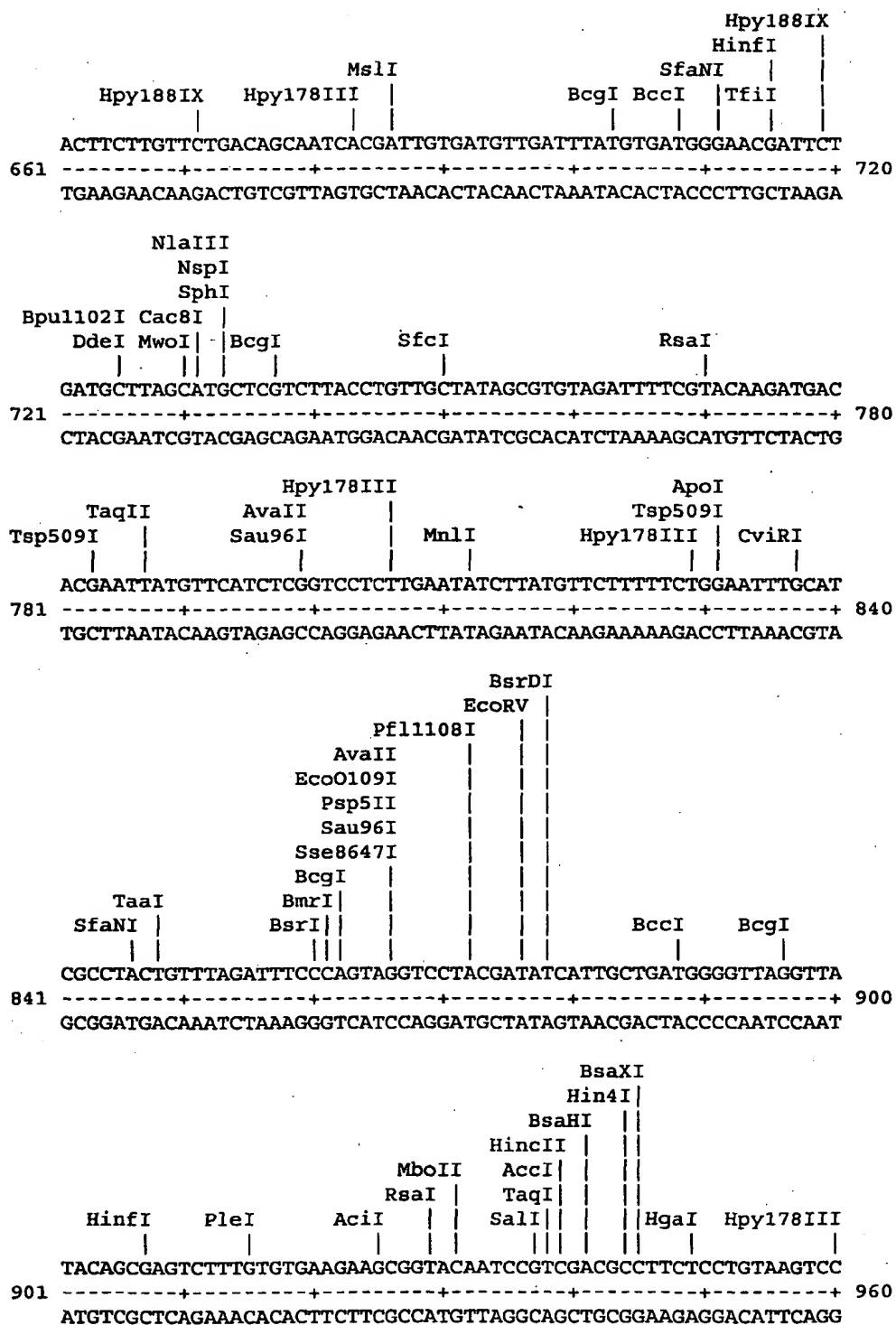


Fig. 10 (con't)

**SUBSTITUTE SHEET (RULE 26)**

Figure 11: Identification of T- and B- cell epitopes from open reading frame ORF: cpn100634 (SEQ ID NO: 11)

```

1 MKTSIPWVLV SSVLAFSCHL QSLANEELLS PDDSFNGNID SGTFTPKTS
51 TTYSLTGDVF FYEPGKGTP L SDSCFKQTTD NLTFGLNGHS LTFGFIDAGT
101 HAGAAASTTA NKNLTFSGFS LLSFDSSPST TVTTGQGTLS SAGGVNLENI
151 RKLVVAGNFS TADGGAIKGA SFLLTGTSGD ALFSNNSSST KGGAIATTAG
201 ARIANNTGYV RFLSNIAS TS GGAIDDEGTS ILSNNKFLYF EGNAAKTTGG
251 AICNTKASGS PELIISNNKT LIFASNVAET SGGAIHAKKL ALSSGGFTEF
301 LRNNVSSATP KGGAISIDAS GELSLSAETG NITFVRNTLT TTGSTDTPKR
351 NAINIGSNGK FTELRAAKNH TIFYDPI TS EGTSSDVLKI NNGSAGALNP
401 YQGTILFSGE TLTADELKVA DNLKSSFTQP VLSGGKLLL QKGVTTLEST
451 FSQEAGSLLG MDSGTTLSTT AGSITITNLG INVDSLGLKQ PVSLTAKGAS
501 NKVIVSGKLN LIDIEGNIYE SHMFSHDQLF SLLKITVDAD VDTNVDISSL
551 IPVPAEDPNS EYGFQGWNV NWTDTATNT KEATATWTKT GFVPSPERKS
601 ALVCNTLWGV FTDIRSLQQL VEIGATGMEH KQGFVSSMT NFLHKTGDEN
651 RKGFRHTSGG YVIGGSAHTP KDDLFTFAFC HLFARDKDCF IAHNNSRTYG
701 GTLFFKHSHT LQPQNYLRLG RAKFSESAIE KFPREIPLAL DVQVSFSHSD
751 NRMETHYTS L PESEGSWSNE CIAGGIGLDL PFVLSNPHPL FKTFIPQMKV
801 EMVYVSQNSF FESSSDGRGF SIGRLLNLSI PVGAKFVQGD IGDSYTYDLS
851 GFFVSDVYRN NPQSTATLVM SPDSWKIRGG NLSRQAFLLR GSNNYVYNSN
901 CELFGHYAME LRGSSRYNV DVGTKLRF

```

Possible T cell epitope:

824 RLLNLSIPV

Possible B cell epitope:

644 HKTGDENRKGFRTS



74/77

Figure 12: Identification of T- and B- cell epitopes from open reading frame ORF: cpn100635 (SEQ ID NO: 12)

```
1  MKSQFSWLVL  SSTLACFTSC  STVFAATAEN  IGPSDSFDGS  TNTGTYTPKN
51  TTTGIDYTLT  GDITLQNLGD  SAALTKGCFs  DTTESLSFAG  KGYSLSFLNI
101 KSSAEGAALS  VTTDKNLSLT  GFSSLTFLAA  PSSVITTPSG  KGAVKCGGDL
151 TFDNNGTILF  KQDYCEENG  AISTKNLSLK  NSTGSISFEG  NKSSATGKKG
201 GAICATGTVD  ITNNTAPTLF  SNNIAEAAG  AINSTGNCTI  TGNTSLVFSE
251 NSVTATAGNG  GALSGDADVT  ISGNQSVTFS  GNQAVANGGA  IYAKKLTLAS
301 GGGGGNPFSN  NIVQGTTAGN  GGAISILAAG  ECSLFSEAGD  HYLNGNAIVA
351 TTPQTTKRNS  IDIGSTGKDH  ELRAISGHSI  FFYDPITANT  AADSTDTLNL
401 NKADAGNSTD  YSGSIVFSGE  KLSEDEAKVA  DNLTSTLKQP  VTLTAGNLVL
451 KRGVTLDTKG  FTQTAGSSVI  MDAGTTLKAS  TEEVTLTGLS  IPVDSLGEKG
501 KVVIAASAAS  KVALSGPIL  LLDNQGNAYE  NHDLGKTQDF  SFVQLSALGT
551 ATTTDVPAVP  TVATPTHYGY  QGTWGMTWVD  DTASTPKTKT  ATLAWTNTGY
601 LPNPERQGPL  VPNSLWGSFS  DIQAIQGVIE  RSALTLCSDR  GFWAAGVANF
651 LDKDKKGEKR  KYRHKSGGYA  IGGAAQTCSE  NLISFAFCQL  FGSDKDFLVA
701 KNHTDTYAGA  FYIQHITECS  GFIGCLLDKL  PGWSHKPLV  LEGQLAYSHV
751 SNDLKTKYTA  YPEVKGSWGN  NAFNMMLGAS  SHSYPEYLHC  FDTYAPYIKL
801 NLTYIRQDSF  SEKGTEGRSF  DDSNLFNLSL  PIGVKFEKFS  DCNDFSYDLT
851 LSYVPDLIRN  DPKCTTALVI  SGASWETYN  NLARQALQVR  AGSHYAFSPM
901 FEVLGQFVFE  VRGSSRIYNV  DLGGKFQF
```

Possible T cell epitope:

903 VLGQFVFEV

Possible B cell epitope:

DKDKKGEKRKYRHKSG

75/77

Figure 13: Identification of T- and B- cell epitopes from open reading frame ORF: cpn100638 (SEQ ID NO: 14)

```

1 MKSSLHWFLI SSSLALPLSL NFSAAFAVVE INLGPTNSFS GPGTYTPPAQ
51 TTNADGTIYN LTGDVSITNA GSPTALTASC FKETTGNLSF QGHGYQFLLO
101 NIDAGANCTF TANTAANKLLS FSGFSYLSLI QTTNATTGTG AIKSTGACSI
151 QSNYSCYFGQ NFSNDNGGAL QGSSISLSLN PNLTFAKNKA TQKGGALYST
201 GGITINNTLN SASFSENTAA NNGGAIYTEA SSFISSNKAI SFINNSVTAT
251 SATGGAIYCS STSAPKPVLT LSDNGELNFI GNTAITSGGA IYTDNLVLSS
301 GGPTLEFKNS GYDTAAPLGG AIAIADSGSL SLSALGGDIT FEGNTVVKGA
351 SSSQTTTRNS INIGNTNAKI VQLRASQGNT IYFYDPITTS ITAALSDALN
401 LNGPDLAGNP AYQGTIVFSG EKLSEAEAAE ADNLKSTIQQ PLTLAGGQLS
451 LKSGVTLVAK SFSQSPGSTL LMDAGTTLET ADGITINNLV LNVDLKETK
501 KGTLKATQAS QTVTLSGSL LVDPSGNVYE DVSWNNPQVF SCLTLTADDP
551 ANIHITDLAA DPLEKNPIHW GYQGNWALSW QEDTATKSKA ATLTWTKTGY
601 NPNPERRGTL VANTLWGSFV DVRSIQQLVA TKVRQSQETR GIWCEGISNF
651 FHKDSTKINK GFRHISAGYV VGATTTLASD NLITAAFCQL FGKDRDHFIN
701 KNRASAYAAS LHLQHLATLS SPSLLRYLPG SESEQPVLFD AQISYIYSKN
751 TMKTYYTQAP KGESSWYNDG CALELASSLP HTALSHEGLF HAYFPFIKVE
801 ASYIHQDSFK ERNTTLVRSF DSGDLINVSV PIGITFERFS RNERASYEAT
851 VIYVADVYRK NPDCTTALLI NNTSWKTTGT NLSRQAGIGR AGIFYAFSPN
901 LEVTSNLSME IRGSSRSYNA DLGGKFQF

```

Possible T cell epitope:

614 TLWGSFVDV

Possible B cell epitope:

595 WTKTGYNPNPERRG

76/77

Figure 14: Identification of T- and B- cell epitopes from open reading frame ORF:  
cpn100639 (SEQ ID NO: 15)

```

1 MKIPLHKLLI SSTLVTPILL SIATYGADAS LSPTDSFDGA GGSTFTPKST
51 ADANGTNYVL SGNVYINDAG KGTALTGCCF TETTGDLTFT GKGYSFSFNT
101 VDAGSNAGAA ASTTADKALI FTGFSNLSFI AAPGTTVASG KSTLSSAGAL
151 NLTDNGTILF SQNVSNNEANN NGGAITTKTL SISGNTSSIT FTSNSAKKLG
201 GAIYSSAAAS ISGNTGQLVF MNNKGETGGG ALGFEASSI TQNSSLFFSG
251 NTATDAAGKG GAIYCEKTGE TPTLTISGNK SLTFAENSSV TQGGAICAHG
301 LDLSAAGPTL FSNNRCGNTA AGKGGAIAIA DSGSLSLSAN QGDITFLGNT
351 LTSTSAPTST RNAIYLGSSA KITNLRAAQG QSIYFYDPIA SNTTGASDVL
401 TINQPDNSNP LDYSGTIVFS GEKLSADEAK AADNFTSILK QPLALASGTL
451 ALKGNVELDV NGFTQTEGST LLMQPGTKLK ADTEAISLTK LVVDLSALEG
501 NKSVSJETAG ANKTITLTSP LVFQDSSGNF YESHTINQAF TQPLVVFTAA
551 TAASDIYIDA LLTSPVQTPE PHYGYQGHWE ATWADTSTAK SGTMTWVTTG
601 YNPNPERRAS VVPDSLWASF TDIRTLQQIM TSQANSIYQQ RGLWASGTAN
651 FFHKDKSGTN QAFRHKSYGY IVGGS AEDFS ENIFSVAFQC LFGKDKDLFI
701 VENTSHNYLA SLYLQHRAFL GGLPMPSPFGS ITDMLKDIPL ILNAQLSYSY
751 TKNDMDTRYT SYPEAQGSWT NNSGALELGG SLALYLPKEA PFFQGYFPFL
801 KFQAVYSRQQ NFKESGAEAR AFDDGDLVNC SIPVGIRLEK ISEDEKNNFE
851 ISLAYIGDVY RKNPRSRTSL MVSGASWTSL CKNLARQAFL ASAGSHLTLS
901 PHVELSGEAA YELRGSAGHIY NVDCGLRYSF

```

Possible T cell epitope:

7 KLLISSTLV

Possible B cell epitopes:

839 EKISEDEKNNF

860 YRKNPRSRT

77/77

Figure 15: Identification of T- and B- cell epitopes from open reading frame ORF: cpn100708 (SEQ ID NO: 16)

```
1  MLSSLIRDSF  PLLILLPTFL  AALGASVAGG  VMGTIYVVKR  IVSISGSISH
51  AILGGIGLTL  WIQYKLHLSF  FPMYGAIVGA  IFLALCIGKI  HLKYQEREDS
101 LIAMIWSVGM  AIGIIFISRL  PTFNGELINF  LFGNILWVTP  SDLYSLGIFD
151 LLVLGIVVLC  HTRFLALCFD  ERYTALNHCS  VQLWYFLLLV  LTAITIVMLI
201 YVMGTILMLS  MLVLPVAIAC  RFSYKMTRIM  FISVLLNILC  SFSGICIAYC
251 LDFPVGPTIS  LLMGLGYTAS  LCVKKRYNPS  TPSPVSPEIN  TNV
```

Possible T cell epitope:

130 FLFGNILWV

Possible B cell epitope:

91 HLKYQEREDS

```
<110> Connaught Laboratories Limited
<120> Chlamydia antigens and corresponding DNA fragments and uses thereof
<130> 77813-2
<140>
<141>
<150> US 60/110,427
<150> US 60/110,438
<150> US 60/110,339
<150> US 60/110,428
<150> US 60/110,340
<151> December 1, 1998
<160> 16
<170> PatentIn Ver. 2.0
```

```
<210> 1
<211> 2950
<212> DNA
<213> Chlamydia pneumoniae
<220>
<221> CDS
<222> (101)..(2884)
<400> 1
```

1

|   |      |
|---|------|
| gca ggt gct gct gca tct aca aca gca aat aag aat ctt acc ttc tca | 451  |
| Ala Gly Ala Ala Ala Ser Thr Thr Ala Asn Lys Asn Leu Thr Phe Ser |      |
| 105 110 115   |      |
| ggg ttt tcc tta ctg agt ttt gat tcc tct cct agc aca acg gtt act | 499  |
| Gly Phe Ser Leu Leu Ser Phe Asp Ser Ser Pro Ser Thr Thr Val Thr |      |
| 120 125 130   |      |
| aca ggt cag gga acg ctt tcc tca gca gga ggc gta aat tta gaa aat | 547  |
| Thr Gly Gln Gly Thr Leu Ser Ser Ala Gly Gly Val Asn Leu Glu Asn |      |
| 135 140 145   |      |
| att cgt aaa ctt gta gtt gct ggg aat ttt tct act gca gat ggt gga | 595  |
| Ile Arg Lys Leu Val Ala Gly Asn Phe Ser Thr Ala Asp Gly Gly     |      |
| 150 155 160 165   |      |
| gct atc aaa gga gcg tct ttc ctt tta act ggc act tct gga gat gct | 643  |
| Ala Ile Lys Gly Ala Ser Phe Leu Leu Thr Gly Thr Ser Gly Asp Ala |      |
| 170 175 180   |      |
| ctt ttt agt aac aac tct tca tca aca aag gga gga gca att gct act | 691  |
| Leu Phe Ser Asn Asn Ser Ser Ser Thr Lys Gly Gly Ala Ile Ala Thr |      |
| 185 190 195   |      |
| aca gca ggc gct cgc ata gca aat aac aca ggt tat gtt aga ttc cta | 739  |
| Thr Ala Gly Ala Arg Ile Ala Asn Asn Thr Gly Tyr Val Arg Phe Leu |      |
| 200 205 210   |      |
| tct aac ata gcg tct acg tca gga ggc gct atc gat gat gaa ggc acg | 787  |
| Ser Asn Ile Ala Ser Thr Ser Gly Gly Ala Ile Asp Asp Glu Gly Thr |      |
| 215 220 225   |      |
| tcg ata cta tcg aac aac aaa ttt cta tat ttt gaa ggc aat gca gcg | 835  |
| Ser Ile Leu Ser Asn Asn Lys Phe Leu Tyr Phe Glu Gly Asn Ala Ala |      |
| 230 235 240 245   |      |
| aaa act act ggc ggt gcg atc tgc aac acc aag gcg agt gga tct cct | 883  |
| Lys Thr Thr Gly Gly Ala Ile Cys Asn Thr Lys Ala Ser Gly Ser Pro |      |
| 250 255 260   |      |
| gaa ctg ata atc tct aac aat aag act ctg atc ttt gct tca aac gta | 931  |
| Glu Leu Ile Ile Ser Asn Asn Lys Thr Leu Ile Phe Ala Ser Asn Val |      |
| 265 270 275   |      |
| gca gaa aca agc ggt ggc gcc atc cat gct aaa aag cta gcc ctt tcc | 979  |
| Ala Glu Thr Ser Gly Gly Ala Ile His Ala Lys Lys Leu Ala Leu Ser |      |
| 280 285 290   |      |
| tct gga ggc ttt aca gag ttt cta cga aat aat gtc tca tca gca act | 1027 |
| Ser Gly Gly Phe Thr Glu Phe Leu Arg Asn Asn Val Ser Ser Ala Thr |      |
| 295 300 305   |      |
| cct aag ggg ggt gct atc agc atc gat gcc tca gga gag ctc agt ctt | 1075 |
| Pro Lys Gly Gly Ala Ile Ser Ile Asp Ala Ser Gly Glu Leu Ser Leu |      |
| 310 315 320 325   |      |

|   |      |
|---|------|
| tct gca gag aca gga aac att acc ttt gta aga aat acc ctt aca aca | 1123 |
| Ser Ala Glu Thr Gly Asn Ile Thr Phe Val Arg Asn Thr Leu Thr Thr |      |
| 330 335 340   |      |
| acc gga agt acc gat act cct aaa cgt aat gcg atc aac ata gga agt | 1171 |
| Thr Gly Ser Thr Asp Thr Pro Lys Arg Asn Ala Ile Asn Ile Gly Ser |      |
| 345 350 355   |      |
| aac ggg aaa ttc acg gaa tta cgg gct gct aaa aat cat aca att ttc | 1219 |
| Asn Gly Lys Phe Thr Glu Leu Arg Ala Ala Lys Asn His Thr Ile Phe |      |
| 360 365 370   |      |
| ttc tat gat ccc atc act tca gaa gga acc tca tca gac gta ttg aag | 1267 |
| Phe Tyr Asp Pro Ile Thr Ser Glu Gly Thr Ser Ser Asp Val Leu Lys |      |
| 375 380 385   |      |
| ata aat aac ggc tct gcg gga gct ctc aat cca tat caa gga acg att | 1315 |
| Ile Asn Asn Gly Ser Ala Gly Ala Leu Asn Pro Tyr Gln Gly Thr Ile |      |
| 390 395 400 405   |      |
| cta ttt tct gga gaa acc cta aca gca gat gaa ctt aaa gtt gct gac | 1363 |
| Leu Phe Ser Gly Glu Thr Leu Thr Ala Asp Glu Leu Lys Val Ala Asp |      |
| 410 415 420   |      |
| aat tta aaa tct tca ttc acg cag cca gtc tcc cta tcc gga gga aag | 1411 |
| Asn Leu Lys Ser Ser Phe Thr Gln Pro Val Ser Leu Ser Gly Gly Lys |      |
| 425 430 435   |      |
| tta ttg cta caa aag gga gtc act tta gag agc acg agc ttc tct caa | 1459 |
| Leu Leu Leu Gln Lys Gly Val Thr Leu Glu Ser Thr Ser Phe Ser Gln |      |
| 440 445 450   |      |
| gag gcc ggt tct ctc ctc ggc atg gat tca gga acg aca tta tca act | 1507 |
| Glu Ala Gly Ser Leu Leu Gly Met Asp Ser Gly Thr Thr Leu Ser Thr |      |
| 455 460 465   |      |
| aca gct ggg agt att aca atc acg aac cta gga atc aat gtt gac tcc | 1555 |
| Thr Ala Gly Ser Ile Thr Ile Thr Asn Leu Gly Ile Asn Val Asp Ser |      |
| 470 475 480 485   |      |
| tta ggt ctt aag cag ccc gtc agc cta aca gca aaa ggt gct tca aat | 1603 |
| Leu Gly Leu Lys Gln Pro Val Ser Leu Thr Ala Lys Gly Ala Ser Asn |      |
| 490 495 500   |      |
| aaa gtg atc gta tct ggg aag ctc aac ctg att gat att gaa ggg aac | 1651 |
| Lys Val Ile Val Ser Gly Lys Leu Asn Leu Ile Asp Ile Glu Gly Asn |      |
| 505 510 515   |      |
| att tat gaa agt cat atg ttc agc cat gac cag ctc ttc tct cta tta | 1699 |
| Ile Tyr Glu Ser His Met Phe Ser His Asp Gln Leu Phe Ser Leu Leu |      |
| 520 525 530   |      |
| aaa atc acg gtt gat gct gat gtt gat act aac gtt gac atc agc agc | 1747 |
| Lys Ile Thr Val Asp Ala Asp Val Asp Thr Asn Val Asp Ile Ser Ser |      |
| 535 540 545   |      |

|   |      |
|---|------|
| ctt atc cct gtt cct gct gag gat cct aat tca gaa tac gga ttc caa | 1795 |
| Leu Ile Pro Val Pro Ala Glu Asp Pro Asn Ser Glu Tyr Gly Phe Gln |      |
| 550 555 560 565   |      |
| gga caa tgg aat gtt aat tgg act acg gat aca gct aca aat aca aaa | 1843 |
| Gly Gln Trp Asn Val Asn Trp Thr Thr Asp Thr Ala Thr Asn Thr Lys |      |
| 570 575 580   |      |
| gag gcc acg gca act tgg acc aaa aca gga ttt gtt ccc agc ccc gaa | 1891 |
| Glu Ala Thr Ala Thr Trp Thr Lys Thr Gly Phe Val Pro Ser Pro Glu |      |
| 585 590 595   |      |
| aga aaa tct gcg tta gta tgc aat acc cta tgg gga gtc ttt act gac | 1939 |
| Arg Lys Ser Ala Leu Val Cys Asn Thr Leu Trp Gly Val Phe Thr Asp |      |
| 600 605 610   |      |
| att cgc tct ctg caa cag ctt gta gag atc ggc gca act ggt atg gaa | 1987 |
| Ile Arg Ser Leu Gln Gln Leu Val Glu Ile Gly Ala Thr Gly Met Glu |      |
| 615 620 625   |      |
| cac aaa caa ggt ttc tgg gtt tcc tcc atg acg aac ttc ctg cat aag | 2035 |
| His Lys Gln Gly Phe Trp Val Ser Ser Met Thr Asn Phe Leu His Lys |      |
| 630 635 640 645   |      |
| act gga gat gaa aat cgc aaa ggc ttc cgt cat acc tct gga ggc tac | 2083 |
| Thr Gly Asp Glu Asn Arg Lys Gly Phe Arg His Thr Ser Gly Gly Tyr |      |
| 650 655 660   |      |
| gtc atc ggt gga agt gct cac act cct aaa gac gac cta ttt acc ttt | 2131 |
| Val Ile Gly Gly Ser Ala His Thr Pro Lys Asp Asp Leu Phe Thr Phe |      |
| 665 670 675   |      |
| gcg ttc tgc cat ctc ttt gct aga gac aaa gat tgt ttt atc gct cac | 2179 |
| Ala Phe Cys His Leu Phe Ala Arg Asp Lys Asp Cys Phe Ile Ala His |      |
| 680 685 690   |      |
| aac aac tct aga acc tac ggt gga act tta ttc ttc aag cac tct cat | 2227 |
| Asn Asn Ser Arg Thr Tyr Gly Gly Thr Leu Phe Phe Lys His Ser His |      |
| 695 700 705   |      |
| acc cta caa ccc caa aac tat ttg aga tta gga aga gca aag ttt tct | 2275 |
| Thr Leu Gln Pro Gln Asn Tyr Leu Arg Leu Gly Arg Ala Lys Phe Ser |      |
| 710 715 720 725   |      |
| gaa tca gct ata gaa aaa ttc cct agg gaa att ccc cta gcc ttg gat | 2323 |
| Glu Ser Ala Ile Glu Lys Phe Pro Arg Glu Ile Pro Leu Ala Leu Asp |      |
| 730 735 740   |      |
| gtc caa gtt tcg ttc agc cat tca gac aac cgt atg gaa acg cac tat | 2371 |
| Val Gln Val Ser Phe Ser His Ser Asp Asn Arg Met Glu Thr His Tyr |      |
| 745 750 755   |      |
| acc tca ttg cca gaa tcc gaa ggt tct tgg agc aac gag tgt ata gct | 2419 |
| Thr Ser Leu Pro Glu Ser Glu Gly Ser Trp Ser Asn Glu Cys Ile Ala |      |
| 760 765 770   |      |



```

ggt ggt atc ggc cta gac ctt cct ttt gtt ctt tcc aac cca cat cct 2467
Gly Gly Ile Gly Leu Asp Leu Pro Phe Val Leu Ser Asn Pro His Pro
775 780 785

ctt ttc aag acc ttc att cca cag atg aaa gtc gaa atg gtt tat gta 2515
Leu Phe Lys Thr Phe Ile Pro Gln Met Lys Val Glu Met Val Tyr Val
790 795 800 805

tca caa aat agc ttc ttc gaa agc tct agt gat ggc cgt ggt ttt agt 2563
Ser Gln Asn Ser Phe Phe Glu Ser Ser Ser Asp Gly Arg Gly Phe Ser
810 815 820

att gga agg ctg ctt aac ctc tcg att cct gtg ggt gcg aaa ttc gtg 2611
Ile Gly Arg Leu Leu Asn Leu Ser Ile Pro Val Gly Ala Lys Phe Val
825 830 835

cag ggg gat atc gga gat tcc tac acc tat gat ctc tca gga ttc ttt 2659
Gln Gly Asp Ile Gly Asp Ser Tyr Thr Tyr Asp Leu Ser Gly Phe Phe
840 845 850

gtt tcc gat gtc tat cgt aac aat ccc caa tct aca gcg act ctt gtg 2707
Val Ser Asp Val Tyr Arg Asn Asn Pro Gln Ser Thr Ala Thr Leu Val
855 860 865

atg agc cca gac tct tgg aaa att cgc ggt ggc aat ctt tca aga cag 2755
Met Ser Pro Asp Ser Trp Lys Ile Arg Gly Gly Asn Leu Ser Arg Gln
870 875 880 885

gca ttt tta ctg agg ggt agc aac aac tac gtc tac aac tcc aat tgt 2803
Ala Phe Leu Leu Arg Gly Ser Asn Asn Tyr Val Tyr Asn Ser Asn Cys
890 895 900

gag ctc ttc gga cat tac gct atg gaa ctc cgt gga tct tca agg aac 2851
Glu Leu Phe Gly His Tyr Ala Met Glu Leu Arg Gly Ser Ser Arg Asn
905 910 915

tac aat gta gat gtt ggt acc aaa ctc cga ttc tagattgcta aaactcccta 2904
Tyr Asn Val Asp Val Gly Thr Lys Leu Arg Phe
920 925

gttctttctag ggagttttct catactttta gggaaatatt tgctat 2950

```

&lt;210&gt; 2

&lt;211&gt; 2784

&lt;212&gt; DNA

&lt;213&gt; Chlamydia pneumoniae

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(2784)

&lt;400&gt; 2

```

atgaagactt cgattccttg ggttttagtt tcctccgtgt tagctttctc atgtcaccta 60
cagtcactag ctaacgagga acttttatca cctgatgata gctttaatgg aaatatcgat 120
tcaggaacgt ttactccaaa aacttcagcc acaacatatt ctctaacagg agatgtcttc 180

```

|             |             |            |            |            |            |      |
|-------------|-------------|------------|------------|------------|------------|------|
| ttttacgagc  | ctggaaaagg  | cactccctta | tctgacagtt | gttttaagca | aaccacggac | 240  |
| aatcttacct  | tcttggggaa  | cggtcatagc | ttaacgtttg | gctttataga | tgctggcact | 300  |
| catgcaggtg  | ctgctgcata  | tacaacagca | aataagaatc | ttaccttctc | agggttttcc | 360  |
| ttactgagtt  | ttgattcctc  | tcctagcaca | acggttacta | caggtcaggg | aacgctttcc | 420  |
| tcagcaggag  | gcgtaaattt  | agaaaatatt | cgtaaacttg | tagttgctgg | gaatttttca | 480  |
| ctgcagatgg  | tggagctatc  | aaaggagcgt | ctttccttta | actggcactt | ctggagatgc | 540  |
| tctttttagt  | aacaactctt  | catcaacaaa | gggaggagca | attgctacta | cagcaggcgc | 600  |
| tcgcatagca  | aataacacag  | gttatgttag | attcctatct | aacatagcgt | ctacgtcagg | 660  |
| aggcgctatc  | gatgatgaag  | gcacgtcgat | actatcgaa  | aacaaatttc | tatatattga | 720  |
| agggaactca  | tcgaaaaacta | ctggcggtgc | gatctgcaac | accaaggcga | gtggatctcc | 780  |
| tgaactgata  | atctctaaca  | ataagactct | gatctttgct | tcaaactgag | cagaaacaag | 840  |
| cggtggcgcc  | atccatgcta  | aaaagctagc | cctttcctct | ggaggcttta | cagagtttct | 900  |
| acgaaataat  | gtctcatcag  | caactcctaa | ggggggtgct | atcagcatcg | atgcctcagg | 960  |
| agagctcaag  | ctttctgcag  | agacaggaaa | cattaccttt | gtaagaaata | cccttacaac | 1020 |
| aaccggaagt  | accgatactc  | ctaaacgtaa | tgcatcaaac | ataggaagta | acgggaaatt | 1080 |
| cacggaatta  | cggtctgcta  | aaaatcatac | aattttcttc | tatgatccca | tcacttcaga | 1140 |
| aggaaactca  | tcagacgtat  | tgaagataaa | taacggctct | gcgggagctc | tcaatccata | 1200 |
| tcaaggaaag  | attctatttt  | ctggagaaac | cctaacagca | gatgaactta | aagttgctga | 1260 |
| caatttaaaa  | tcttcattca  | cgcagccagt | ctccctatcc | ggaggaaagt | tattgctaca | 1320 |
| aaaggaggtc  | acttttagaga | gcacgagctt | ctctcaagag | gccggttctc | tcctcggcat | 1380 |
| ggattcagga  | acgacattat  | caactacagc | tgggagtatt | acaatcacga | acctaggaat | 1440 |
| caatgttgac  | tccttaggtc  | ttaagcagcc | cgtcagccta | acagcaaaag | gtgcttcaaa | 1500 |
| taaaagtgatc | gtatctggga  | agctcaacct | gattgatatt | gaagggaaca | tttatgaaag | 1560 |
| tcataatgttc | agccatgacc  | agctcttctc | tctattaaaa | atcacggttg | atgctgatgt | 1620 |
| tgatactaac  | gttgacatca  | gcagccttat | ccctgttcct | gctgaggatc | ctaattcaga | 1680 |
| atacggtatc  | caaggacaat  | ggaatgttaa | ttggactacg | gatacagcta | caaatacaaa | 1740 |
| agaggccacg  | gcaacttgga  | ccaaaacagg | atgtgttccc | agccccgaaa | gaaaatctgc | 1800 |
| gtagtatgac  | aataccctat  | ggggagtctt | tactgacatt | cgctctctgc | aacagcttgt | 1860 |
| agagatcggc  | gcaactggta  | tggaaacaaa | acaaggtttc | tgggtttcct | ccatgacgaa | 1920 |
| cttcctgcat  | aagactggag  | atgaaaatcg | caaaggcttc | cgtcatacct | ctggaggcta | 1980 |
| cgtcatcggt  | ggaagtgttc  | acactcctaa | agacgacct  | tttaccttg  | cgttctgcca | 2040 |
| tctctttgct  | agagacaaa   | attgttttat | atcgctcaca | acaactctag | aacctacggt | 2100 |
| ggaactttat  | tcttcaagca  | ctctcatacc | ctacaacccc | aaaactattt | gagattagga | 2160 |
| agagcaaaag  | tttctgaatc  | agctatagaa | aaattcccta | gggaaattcc | cctagccttg | 2220 |
| gatgtccaag  | tttcgttcag  | ccattcagac | aaccgtatgg | aaacgcacta | tacctcattg | 2280 |
| ccagaatccg  | aaggttcttg  | gagcaacgag | tgtatagctg | gtggtatcgg | cctagacctt | 2340 |
| ccttttgctc  | tttccaaccc  | acatcctctt | ttcaagacct | tcattccaca | gatgaaagtc | 2400 |
| gaaatggttt  | atgtatcaca  | aaatagcttc | ttcgaaagct | ctagtgatgg | ccgtgggttt | 2460 |
| agtattggaa  | ggctgcttaa  | cctctcgatt | cctgtgggtg | cgaaattcgt | gcagggggat | 2520 |
| atcgagatt   | cctacacct   | tgatctctca | ggattctttg | tttccgatgt | ctatcgtaac | 2580 |
| aatccccaat  | ctacagcgac  | tcttgtgatg | agcccagact | cttggaatat | tcgcggtggc | 2640 |
| aatctttcaa  | gacaggcatt  | tttactgagg | ggtagcaaca | actacgtcta | caactccaat | 2700 |
| tgtgagctct  | tcggacatta  | cgctatggaa | ctccgtggat | cttcaaggaa | ctacaatgta | 2760 |
| gatgttggtg  | ccaaactccg  | attc       |            |            |            | 2784 |

<210> 3  
 <211> 2950  
 <212> DNA  
 <213> Chlamydia pneumoniae  
 <220>  
 <221> CDS  
 <222> (101)..(2884)  
 <400> 3

7

|   |      |
|---|------|
| aat acg gct cct acc ctc ttc tcg aac aat att gct gaa gct gca ggt | 787  |
| Asn Thr Ala Pro Thr Leu Phe Ser Asn Asn Ile Ala Glu Ala Ala Gly |      |
| 215 220 225   |      |
| gga gct ata aat agc aca gga aac tgt aca att aca ggg aat acg tct | 835  |
| Gly Ala Ile Asn Ser Thr Gly Asn Cys Thr Ile Thr Gly Asn Thr Ser |      |
| 230 235 240 245   |      |
| ctt gta ttt tct gaa aat agt gtg aca gcg acc gca gga aat gga gga | 883  |
| Leu Val Phe Ser Glu Asn Ser Val Thr Ala Thr Ala Gly Asn Gly Gly |      |
| 250 255 260   |      |
| gct ctt tct gga gat gcc gat gtt acc ata tct ggg aat cag agt gta | 931  |
| Ala Leu Ser Gly Asp Ala Asp Val Thr Ile Ser Gly Asn Gln Ser Val |      |
| 265 270 275   |      |
| act ttc tca gga aac caa gct gta gct aat ggc gga gcc att tat gct | 979  |
| Thr Phe Ser Gly Asn Gln Ala Val Ala Asn Gly Gly Ala Ile Tyr Ala |      |
| 280 285 290   |      |
| aag aag ctt aca ctg gct tcc ggg ggg ggg ggg ggg aat ccc ttt tct | 1027 |
| Lys Lys Leu Thr Leu Ala Ser Gly Gly Gly Gly Gly Asn Pro Phe Ser |      |
| 295 300 305   |      |
| aac aat ata gtc caa ggt acc act gca ggt aat ggt gga gcc att tct | 1075 |
| Asn Asn Ile Val Gln Gly Thr Thr Ala Gly Asn Gly Gly Ala Ile Ser |      |
| 310 315 320 325   |      |
| ata ctg gca gct gga gag tgt agt ctt ttc agc gaa gca ggg gac cat | 1123 |
| Ile Leu Ala Ala Gly Glu Cys Ser Leu Phe Ser Glu Ala Gly Asp His |      |
| 330 335 340   |      |
| tac ctt aat ggg aat gcc att gtt gca act aca cca caa act aca aaa | 1171 |
| Tyr Leu Asn Gly Asn Ala Ile Val Ala Thr Thr Pro Gln Thr Thr Lys |      |
| 345 350 355   |      |
| aga aat tct att gac ata gga tct act ggc aaa gat cac gaa tta cgt | 1219 |
| Arg Asn Ser Ile Asp Ile Gly Ser Thr Gly Lys Asp His Glu Leu Arg |      |
| 360 365 370   |      |
| gca ata tct ggg cat agc atc ttt ttc tac gat ccg att act gct aat | 1267 |
| Ala Ile Ser Gly His Ser Ile Phe Phe Tyr Asp Pro Ile Thr Ala Asn |      |
| 375 380 385   |      |
| acg gct gcg gat tct aca gat act tta aat ctc aat aag gct gat gca | 1315 |
| Thr Ala Ala Asp Ser Thr Asp Thr Leu Asn Leu Asn Lys Ala Asp Ala |      |
| 390 395 400 405   |      |
| ggt aat agt aca gat tat agt ggg tcg att gtt ttt tct ggt gaa aag | 1363 |
| Gly Asn Ser Thr Asp Tyr Ser Gly Ser Ile Val Phe Ser Gly Glu Lys |      |
| 410 415 420   |      |
| ctc tct gaa gat gaa gca aaa gtt gca gac aac ctc act tct acg ctg | 1411 |
| Leu Ser Glu Asp Glu Ala Lys Val Ala Asp Asn Leu Thr Ser Thr Leu |      |
| 425 430 435   |      |

|   |      |
|---|------|
| aag cag cct gta act cta act gca gga aat tta gta ctt aaa cgt ggt | 1459 |
| Lys Gln Pro Val Thr Leu Thr Ala Gly Asn Leu Val Leu Lys Arg Gly |      |
| 440 445 450   |      |
| gtc act ctc gat acg aaa ggc ttt act cag acc gcg ggt tcc tct gtt | 1507 |
| Val Thr Leu Asp Thr Lys Gly Phe Thr Gln Thr Ala Gly Ser Ser Val |      |
| 455 460 465   |      |
| att atg gat gcg ggc aca acg tta aaa gca agt aca gag gag gtc act | 1555 |
| Ile Met Asp Ala Gly Thr Thr Leu Lys Ala Ser Thr Glu Glu Val Thr |      |
| 470 475 480 485   |      |
| tta aca ggt ctt tcc att cct gta gac tct tta ggc gag ggt aag aaa | 1603 |
| Leu Thr Gly Leu Ser Ile Pro Val Asp Ser Leu Gly Glu Gly Lys Lys |      |
| 490 495 500   |      |
| gtt gta att gct gct tct gca gca agt aaa aat gta gcc ctt agt ggt | 1651 |
| Val Val Ile Ala Ala Ser Ala Ala Ser Lys Asn Val Ala Leu Ser Gly |      |
| 505 510 515   |      |
| ccg att ctt ctt ttg gat aac caa ggg aat gct tat gaa aat cac gac | 1699 |
| Pro Ile Leu Leu Leu Asp Asn Gln Gly Asn Ala Tyr Glu Asn His Asp |      |
| 520 525 530   |      |
| tta gga aaa act caa gac ttt tca ttt gtg cag ctc tct gct ctg ggt | 1747 |
| Leu Gly Lys Thr Gln Asp Phe Ser Phe Val Gln Leu Ser Ala Leu Gly |      |
| 535 540 545   |      |
| act gca aca act aca gat gtt cca gcg gtt cct aca gta gca act cct | 1795 |
| Thr Ala Thr Thr Thr Asp Val Pro Ala Val Pro Thr Val Ala Thr Pro |      |
| 550 555 560 565   |      |
| acg cac tat ggg tat caa ggt act tgg gga atg act tgg gtt gat gat | 1843 |
| Thr His Tyr Gly Tyr Gln Gly Thr Trp Gly Met Thr Trp Val Asp Asp |      |
| 570 575 580   |      |
| acc gca agc act cca aag act aag aca gcg aca tta gct tgg acc aat | 1891 |
| Thr Ala Ser Thr Pro Lys Thr Lys Thr Ala Thr Leu Ala Trp Thr Asn |      |
| 585 590 595   |      |
| aca ggc tac ctt ccg aat cct gag cgt caa gga cct tta gtt cct aat | 1939 |
| Thr Gly Tyr Leu Pro Asn Pro Glu Arg Gln Gly Pro Leu Val Pro Asn |      |
| 600 605 610   |      |
| agc ctt tgg gga tct ttt tca gac atc caa gcg att caa ggt gtc ata | 1987 |
| Ser Leu Trp Gly Ser Phe Ser Asp Ile Gln Ala Ile Gln Gly Val Ile |      |
| 615 620 625   |      |
| gag aga agt gct ttg act ctt tgt tca gat cga ggc ttc tgg gct gcg | 2035 |
| Glu Arg Ser Ala Leu Thr Leu Cys Ser Asp Arg Gly Phe Trp Ala Ala |      |
| 630 635 640 645   |      |
| gga gtc gcc aat ttc tta gat aaa gat aag aaa ggg gaa aaa cgc aaa | 2083 |
| Gly Val Ala Asn Phe Leu Asp Lys Asp Lys Lys Gly Glu Lys Arg Lys |      |
| 650 655 660   |      |

|   |      |
|---|------|
| tac cgt cat aaa tct ggt gga tat gct atc gga ggt gca gcg caa act | 2131 |
| Tyr Arg His Lys Ser Gly Gly Tyr Ala Ile Gly Gly Ala Ala Gln Thr |      |
| 665 670 675   |      |
| tgt tct gaa aac tta att agc ttt gcc ttt tgc caa ctc ttt ggt agc | 2179 |
| Cys Ser Glu Asn Leu Ile Ser Phe Ala Phe Cys Gln Leu Phe Gly Ser |      |
| 680 685 690   |      |
| gat aaa gat ttc tta gtc gct aaa aat cat act gat acc tat gca gga | 2227 |
| Asp Lys Asp Phe Leu Val Ala Lys Asn His Thr Asp Thr Tyr Ala Gly |      |
| 695 700 705   |      |
| gcc ttc tat atc caa cac att aca gaa tgt agt ggg ttc ata ggt tgt | 2275 |
| Ala Phe Tyr Ile Gln His Ile Thr Glu Cys Ser Gly Phe Ile Gly Cys |      |
| 710 715 720 725   |      |
| ctc tta gat aaa ctt cct ggc tct tgg agt cat aaa ccc ctc gtt tta | 2323 |
| Leu Leu Asp Lys Leu Pro Gly Ser Trp Ser His Lys Pro Leu Val Leu |      |
| 730 735 740   |      |
| gaa ggg cag ctc gct tat agc cac gtc agt aat gat ctg aag aca aag | 2371 |
| Glu Gly Gln Leu Ala Tyr Ser His Val Ser Asn Asp Leu Lys Thr Lys |      |
| 745 750 755   |      |
| tat act gcg tat cct gag gtg aaa ggt tct tgg ggg aat aat gct ttt | 2419 |
| Tyr Thr Ala Tyr Pro Glu Val Lys Gly Ser Trp Gly Asn Asn Ala Phe |      |
| 760 765 770   |      |
| aac atg atg ttg gga gct tct tct cat tct tat cct gaa tac ctg cat | 2467 |
| Asn Met Met Leu Gly Ala Ser Ser His Ser Tyr Pro Glu Tyr Leu His |      |
| 775 780 785   |      |
| tgt ttt gat acc tat gct cca tac atc aaa ctg aat ctg acc tat ata | 2515 |
| Cys Phe Asp Thr Tyr Ala Pro Tyr Ile Lys Leu Asn Leu Thr Tyr Ile |      |
| 790 795 800 805   |      |
| cgt cag gac agc ttc tcg gag aaa ggt aca gaa gga aga tct ttt gat | 2563 |
| Arg Gln Asp Ser Phe Ser Glu Lys Gly Thr Glu Gly Arg Ser Phe Asp |      |
| 810 815 820   |      |
| gac agc aac ctc ttc aat tta tct ttg cct ata ggg gtg aag ttt gag | 2611 |
| Asp Ser Asn Leu Phe Asn Leu Ser Leu Pro Ile Gly Val Lys Phe Glu |      |
| 825 830 835   |      |
| aag ttc tct gat tgt aat gac ttt tct tat gat ctg act tta tcc tat | 2659 |
| Lys Phe Ser Asp Cys Asn Asp Phe Ser Tyr Asp Leu Thr Leu Ser Tyr |      |
| 840 845 850   |      |
| gtt cct gat ctt atc cgc aat gat ccc aaa tgc act aca gca ctt gta | 2707 |
| Val Pro Asp Leu Ile Arg Asn Asp Pro Lys Cys Thr Thr Ala Leu Val |      |
| 855 860 865   |      |
| atc agc gga gcc tct tgg gaa act tat gcc aat aac tta gca cga cag | 2755 |
| Ile Ser Gly Ala Ser Trp Glu Thr Tyr Ala Asn Asn Leu Ala Arg Gln |      |
| 870 875 880 885   |      |

```

gcc ttg caa gtg cgt gca ggc agt cac tac gcc ttc tct cct atg ttt 2803
Ala Leu Gln Val Arg Ala Gly Ser His Tyr Ala Phe Ser Pro Met Phe
      890                      895                      900

gaa gtg ctc ggc cag ttt gtc ttt gaa gtt cgt gga tcc tca cgg att 2851
Glu Val Leu Gly Gln Phe Val Phe Glu Val Arg Gly Ser Ser Arg Ile
      905                      910                      915

tat aat gta gat ctt ggg ggt aag ttc caa ttc taggagcgtc tctcatgtct 2904
Tyr Asn Val Asp Leu Gly Gly Lys Phe Gln Phe
      920                      925

cagaaattct gagagagatc gcatttagga ttttcttaaa cacgac 2950

```

&lt;210&gt;4

&lt;211&gt;2950

&lt;212&gt;DNA

&lt;213&gt;Chlamydia pneumoniae

&lt;220&gt;

&lt;221&gt;CDS

&lt;222&gt;(101)..(2784)

&lt;400&gt;4

```

atgaaatcgc aattttcctg gttagtgtc tcttcgacat tggcatgttt tactagtgtg 60
tccactgttt ttgctgcaac tgctgaaaat ataggccctt ctgatagctt tgacggaagt 120
actaacacag gcacctatac tcctaaaaat acgactactg gaatagacta tactctgaca 180
ggagatataa ctctgcaaaa ccttggggat tcggcagctt taacgaaggg ttgtttttct 240
gacactacgg aatctttaag ctttgccggt aagggggtact cactttcttt tttaaatatt 300
aagtctagtg ctgaaggcgc agccctttct gttacaactg ataaaaatct gtcgctaaca 360
ggatttttca gtcttacttt cttagcggcc ccatacatcg taatcacaac cccctcagga 420
aaaggtgcag ttaaattgtg aggggatctt acatttgata acaatggaac tatttttatt 480
aaacaagatt actgtgagga aaatggcgga gccatttcta ccaagaatct ttctttgaaa 540
aacagcacgg gatcgatttc ttttgaaggg aataaatcga gcgcaacagg gaaaaaaggt 600
ggggctattt gtgctactgg tactgtagat attacaaata atacggctcc taccctcttc 660
tcgaacaata ttgctgaagc tgcaagtgga gctataaata gcacaggaaa ctgtacaatt 720
acagggaata cgtctcttgt attttctgaa aatagtgtga cagcgaccgc aggaaatgga 780
ggagctcttt ctggagatgc cgatgttacc atactggga atcagagtgt aactttctca 840
ggaaaccaag ctgtagctaa tggcggagcc atttatgcta agaagcttac actggcttcc 900
gggggggggg gggggaatcc cttttctaac aatatagtcc aagggtaccac tgcaggtaat 960
ggtggagcca tttctatact ggcagctgga gagtgtagtc ttttcagcga agcaggggac 1020
cattaccta atgggaatgc cattgttgca actacaccac aaactacaaa aagaaattca 1080
ttgacatagg atctactggc aaagatcacg attacgtgca atatctgggc atagcatctt 1140
tttctacgat ccgattactg ctaatacggc tgcggattct acagatactt taaatctcaa 1200
taaggctgat gcaggttaata gtacagatta tagtgggtcg attgtttttt ctggtgaaaa 1260
gctctctgaa gatgaagcaa aagttgcaga caacctcact tctacgctga agcagcctgt 1320
aactctaact gcaggaaatt tagtacttaa acgtggtgtc actctcgata cgaaaggctt 1380
tactcagacc gcgggttccct ctgttattat ggatgcgggc acaacgttaa aagcaagtac 1440
agaggaggtc actttaacag gtctttccat tcctgtagac tctttaggcg agggtaagaa 1500
agttgtaatt gctgcttctg cagcaagtaa aaatgtagcc cttagtgggtc cgattcttct 1560
tttgataaac caagggaatg cttatgaaaa tcacgactta ggaaaaactc aagacttttc 1620
atttgtgcag ctctctgtc tgggtactgc aacaactaca gatgttccag cggttcctac 1680
agtagcaact cctacgcact atgggtatca aggtacttgg ggaatgactt gggttgatga 1740
taccgcaagc actccaaaga ctaagacagc gacattagct tggaccaatc aggctacctt 1800
ccgaatcctg agcgtcaagg acctttagtt cctaatagcc tttggggatc tttttcagac 1860

```

```

atccaagcga ttcaaggtgt catagagaga agtgctttga ctctttgttc agatcgaggc 1920
ttctgggctg cgggagtcgc caatttctta gataaagata agaaagggga aaaacgcaaa 1980
taccgtcata aatctggttg atatgctatc ggaggtgcag cgcaaacttg ttctgaaaac 2040
ttaattagct ttgccttttg ccaactcttt ggtagcgata aagatttctt agtcgctaaa 2100
aatcatactg atacctatgc aggagccttc tatatccaac acattacaga atgtagtggg 2160
ttcatagggt gtctcttaga taaacttcct ggctcttgga gtcataaacc cctcgtttta 2220
gaagggcagc tcgcttatag ccacgtcagt aatgatctga agacaaagta tactgcgtat 2280
cctgaggtga aaggttcttg ggggaataat gcttttaaca tgatgttggg agcttcttct 2340
cattcttata ctgaatacct gcattgtttt gatacctatg ctccatacat caaactgaat 2400
ctgacctata tacgtcagga cagcttctcg gagaaagtac agaaggaaga tcttttgatg 2460
acagcaacct cttcaattta tctttgccta taggggtgaa gtttgagaag ttctctgatt 2520
gtaatgactt ttcttatgat ctgactttat cctatgttcc tgatcttata cgcaatgata 2580
ccaaatgcac tacagcactt gtaatcagcg gagccactt cttgggaaac ttatgccaat 2640
aacttagcac gacaggcctt gcaagtgcgt gcaggcagtc actacgcctt ctctcctatg 2700
tttgaagtgc tcggccagtt tgtctttgaa gttcgtggat cctcacggat ttataatgta 2760
gatcttgggg gtaagttcca attc 2784

```

```

<210> 5
<211> 2950
<212> DNA
<213> Chlamydia pneumoniae
<220>
<221> CDS
<222> (101)..(2884)
<400> 5

```

```

tgtagattct taacttacta gtctctcctt tcctcttgct ttctttaatt tattgcagta 60

tgtggtgaaa taatttgta aaccacctat agccctctac atg aaa tcc tct ctt 115
                               Met Lys Ser Ser Leu
                               1           5

cat tgg ttt tta atc tcg tca tct tta gca ctt ccc ttg tca cta aat 163
His Trp Phe Leu Ile Ser Ser Ser Leu Ala Leu Pro Leu Ser Leu Asn
          10           15           20

ttc tct gcg ttt gct gct gtt gtt gaa atc aat cta gga cct acc aat 211
Phe Ser Ala Phe Ala Ala Val Val Glu Ile Asn Leu Gly Pro Thr Asn
          25           30           35

agc ttc tct gga cca gga acc tac act cct cca gcc caa aca aca aat 259
Ser Phe Ser Gly Pro Gly Thr Tyr Thr Pro Pro Ala Gln Thr Thr Asn
          40           45           50

gca gat gga act atc tat aat cta aca ggg gat gtc tca atc acc aat 307
Ala Asp Gly Thr Ile Tyr Asn Leu Thr Gly Asp Val Ser Ile Thr Asn
          55           60           65

gca gga tct ccg aca gct cta acc gct tcc tgc ttt aaa gaa act act 355
Ala Gly Ser Pro Thr Ala Leu Thr Ala Ser Cys Phe Lys Glu Thr Thr
          70           75           80           85

```



|   |      |
|---|------|
| ggg aat ctt tct ttc caa ggc cac ggc tac caa ttt ctc cta caa aat | 403  |
| Gly Asn Leu Ser Phe Gln Gly His Gly Tyr Gln Phe Leu Leu Gln Asn |      |
| 90 95 100   |      |
| atc gat gcg gga gcg aac tgt acc ttt acc aat aca gct gca aat aag | 451  |
| Ile Asp Ala Gly Ala Asn Cys Thr Phe Thr Asn Thr Ala Ala Asn Lys |      |
| 105 110 115   |      |
| ctt ctc tcc ttt tca gga ttc tcc tat ttg tca cta ata caa acc acg | 499  |
| Leu Leu Ser Phe Ser Gly Phe Ser Tyr Leu Ser Leu Ile Gln Thr Thr |      |
| 120 125 130   |      |
| aat gct acc aca gga aca gga gcc atc aag tcc aca gga gct tgt tct | 547  |
| Asn Ala Thr Thr Gly Thr Gly Ala Ile Lys Ser Thr Gly Ala Cys Ser |      |
| 135 140 145   |      |
| att cag tcg aac tat agt tgc tac ttt ggc caa aac ttt tct aat gac | 595  |
| Ile Gln Ser Asn Tyr Ser Cys Tyr Phe Gly Gln Asn Phe Ser Asn Asp |      |
| 150 155 160 165   |      |
| aat gga ggc gcc ctc caa ggc agc tct atc agt cta tcg cta aac ccc | 643  |
| Asn Gly Gly Ala Leu Gln Gly Ser Ser Ile Ser Leu Ser Leu Asn Pro |      |
| 170 175 180   |      |
| aac cta acg ttt gcc aaa aac aaa gca acg caa aaa ggg ggt gcc ctc | 691  |
| Asn Leu Thr Phe Ala Lys Asn Lys Ala Thr Gln Lys Gly Gly Ala Leu |      |
| 185 190 195   |      |
| tat tcc acg gga ggg att aca att aac aat acg tta aac tca gca tca | 739  |
| Tyr Ser Thr Gly Gly Ile Thr Ile Asn Asn Thr Leu Asn Ser Ala Ser |      |
| 200 205 210   |      |
| ttt tct gaa aat acc gcg gcg aac aat ggc gga gcc att tac acg gaa | 787  |
| Phe Ser Glu Asn Thr Ala Ala Asn Asn Gly Gly Ala Ile Tyr Thr Glu |      |
| 215 220 225   |      |
| gct agc agt ttt att agc agc aac aaa gca att agc ttt ata aac aat | 835  |
| Ala Ser Ser Phe Ile Ser Ser Asn Lys Ala Ile Ser Phe Ile Asn Asn |      |
| 230 235 240 245   |      |
| agt gtg acc gca acc tca gct aca ggg gga gcc att tac tgt agt agt | 883  |
| Ser Val Thr Ala Thr Ser Ala Thr Gly Gly Ala Ile Tyr Cys Ser Ser |      |
| 250 255 260   |      |
| aca tca gcc ccc aaa cca gtc tta act cta tca gac aac ggg gaa ctg | 931  |
| Thr Ser Ala Pro Lys Pro Val Leu Thr Leu Ser Asp Asn Gly Glu Leu |      |
| 265 270 275   |      |
| aac ttt ata gga aat aca gca att act agt ggt ggg gcg att tat act | 979  |
| Asn Phe Ile Gly Asn Thr Ala Ile Thr Ser Gly Gly Ala Ile Tyr Thr |      |
| 280 285 290   |      |
| gac aat cta gtt ctt tct tct gga gga cct acg ctt ttt aaa aac aac | 1027 |
| Asp Asn Leu Val Leu Ser Ser Gly Gly Pro Thr Leu Phe Lys Asn Asn |      |
| 295 300 305   |      |

|   |      |
|---|------|
| tct ggc tat gat act gca gct ccc tta gga gga gca att gcg att gct | 1075 |
| Ser Gly Tyr Asp Thr Ala Ala Pro Leu Gly Gly Ala Ile Ala Ile Ala |      |
| 310 315 320 325   |      |
| gac tct gga tct ttg agt ctt tcg gct ctt ggt gga gac atc act ttt | 1123 |
| Asp Ser Gly Ser Leu Ser Leu Ser Ala Leu Gly Gly Asp Ile Thr Phe |      |
| 330 335 340   |      |
| gaa gga aac aca gta gtc aaa gga gct tct tcg agt cag acc act acc | 1171 |
| Glu Gly Asn Thr Val Val Lys Gly Ala Ser Ser Ser Gln Thr Thr Thr |      |
| 345 350 355   |      |
| aga aat tct att aac atc gga aac acc aat gct aag att gta cag ctg | 1219 |
| Arg Asn Ser Ile Asn Ile Gly Asn Thr Asn Ala Lys Ile Val Gln Leu |      |
| 360 365 370   |      |
| cga gcc tct caa ggc aat act atc tac ttc tat gat cct ata aca act | 1267 |
| Arg Ala Ser Gln Gly Asn Thr Ile Tyr Phe Tyr Asp Pro Ile Thr Thr |      |
| 375 380 385   |      |
| agc atc act gca gct ctc tca gat gct cta aac tta aat ggt cct gac | 1315 |
| Ser Ile Thr Ala Ala Leu Ser Asp Ala Leu Asn Leu Asn Gly Pro Asp |      |
| 390 395 400 405   |      |
| ctt gca ggg aat cct gca tat caa gga acc atc gta ttt tct gga gag | 1363 |
| Leu Ala Gly Asn Pro Ala Tyr Gln Gly Thr Ile Val Phe Ser Gly Glu |      |
| 410 415 420   |      |
| aag ctc tcg gaa gca gaa gct gca gaa gct gat aat ctc aaa tct aca | 1411 |
| Lys Leu Ser Glu Ala Glu Ala Ala Glu Ala Asp Asn Leu Lys Ser Thr |      |
| 425 430 435   |      |
| att cag caa cct cta act ctt gcg gga ggg caa ctc tct ctt aaa tca | 1459 |
| Ile Gln Gln Pro Leu Thr Leu Ala Gly Gly Gln Leu Ser Leu Lys Ser |      |
| 440 445 450   |      |
| gga gtc act cta gtt gct aag tcc ttt tcg caa tct ccg ggc tct acc | 1507 |
| Gly Val Thr Leu Val Ala Lys Ser Phe Ser Gln Ser Pro Gly Ser Thr |      |
| 455 460 465   |      |
| ctc ctc atg gat gca ggg acc aca tta gaa acc gct gat ggg atc act | 1555 |
| Leu Leu Met Asp Ala Gly Thr Thr Leu Glu Thr Ala Asp Gly Ile Thr |      |
| 470 475 480 485   |      |
| atc aat aat ctt gtt ctc aat gta gat tcc tta aaa gag acc aag aag | 1603 |
| Ile Asn Asn Leu Val Leu Asn Val Asp Ser Leu Lys Glu Thr Lys Lys |      |
| 490 495 500   |      |
| ggc acg cta aaa gca aca caa gca agt cag aca gtc act tta tct gga | 1651 |
| Gly Thr Leu Lys Ala Thr Gln Ala Ser Gln Thr Val Thr Leu Ser Gly |      |
| 505 510 515   |      |
| tcg ctc tct ctt gta gat cct tct gga aat gtc tac gaa gat gtc tct | 1699 |
| Ser Leu Ser Leu Val Asp Pro Ser Gly Asn Val Tyr Glu Asp Val Ser |      |
| 520 525 530   |      |

|   |      |
|---|------|
| tgg aat aac cct caa gtc ttt tct tgt ctc act ctt act gct gac gac | 1747 |
| Trp Asn Asn Pro Gln Val Phe Ser Cys Leu Thr Leu Thr Ala Asp Asp |      |
| 535 540 545   |      |
| ccc gcg aat att cac atc aca gac tta gct gct gat ccc cta gaa aaa | 1795 |
| Pro Ala Asn Ile His Ile Thr Asp Leu Ala Ala Asp Pro Leu Glu Lys |      |
| 550 555 560 565   |      |
| aat cct atc cat tgg gga tac caa ggg aat tgg gca tta tct tgg caa | 1843 |
| Asn Pro Ile His Trp Gly Tyr Gln Gly Asn Trp Ala Leu Ser Trp Gln |      |
| 570 575 580   |      |
| gag gat act gcg act aaa tcc aaa gca gcg act ctt acc tgg aca aaa | 1891 |
| Glu Asp Thr Ala Thr Lys Ser Lys Ala Ala Thr Leu Thr Trp Thr Lys |      |
| 585 590 595   |      |
| aca gga tac aat ccg aat cct gag cgt cgt gga acc tta gtt gct aac | 1939 |
| Thr Gly Tyr Asn Pro Asn Pro Glu Arg Arg Gly Thr Leu Val Ala Asn |      |
| 600 605 610   |      |
| acg cta tgg gga tcc ttt gtt gat gtg cgc tcc ata caa cag ctt gta | 1987 |
| Thr Leu Trp Gly Ser Phe Val Asp Val Arg Ser Ile Gln Gln Leu Val |      |
| 615 620 625   |      |
| gcc act aaa gta cgc caa tct caa gaa act cgc ggc atc tgg tgt gaa | 2035 |
| Ala Thr Lys Val Arg Gln Ser Gln Glu Thr Arg Gly Ile Trp Cys Glu |      |
| 630 635 640 645   |      |
| ggg atc tcg aac ttc ttc cat aaa gat agc acg aag ata aat aaa ggt | 2083 |
| Gly Ile Ser Asn Phe Phe His Lys Asp Ser Thr Lys Ile Asn Lys Gly |      |
| 650 655 660   |      |
| ttt cgc cac ata agt gca ggt tat gtt gta gga gcg act aca aca tta | 2131 |
| Phe Arg His Ile Ser Ala Gly Tyr Val Val Gly Ala Thr Thr Thr Leu |      |
| 665 670 675   |      |
| gct tct gat aat ctt atc act gca gcc ttc tgc caa tta ttc ggg aaa | 2179 |
| Ala Ser Asp Asn Leu Ile Thr Ala Ala Phe Cys Gln Leu Phe Gly Lys |      |
| 680 685 690   |      |
| gat aga gat cac ttt ata aat aaa aat aga gct tct gcc tat gca gct | 2227 |
| Asp Arg Asp His Phe Ile Asn Lys Asn Arg Ala Ser Ala Tyr Ala Ala |      |
| 695 700 705   |      |
| tct ctc cat ctc cag cat cta gcg acc ttg tct tct cca agc ttg tta | 2275 |
| Ser Leu His Leu Gln His Leu Ala Thr Leu Ser Ser Pro Ser Leu Leu |      |
| 710 715 720 725   |      |
| cgc tac ctt cct gga tct gaa agt gag cag cct gtc ctc ttt gat gct | 2323 |
| Arg Tyr Leu Pro Gly Ser Glu Ser Glu Gln Pro Val Leu Phe Asp Ala |      |
| 730 735 740   |      |
| cag atc agc tat atc tat agt aaa aat act atg aaa acc tat tac acc | 2371 |
| Gln Ile Ser Tyr Ile Tyr Ser Lys Asn Thr Met Lys Thr Tyr Tyr Thr |      |
| 745 750 755   |      |

caa gca cca aag gga gag agc tcg tgg tat aat gac ggt tgc gct ctg 2419  
 Gln Ala Pro Lys Gly Glu Ser Ser Trp Tyr Asn Asp Gly Cys Ala Leu  
 760 765 770

gaa ctt gcg agc tcc cta cca cac act gct tta agc cat gag ggt ctc 2467  
 Glu Leu Ala Ser Ser Leu Pro His Thr Ala Leu Ser His Glu Gly Leu  
 775 780 785

ttc cac gcg tat ttt cct ttc atc aaa gta gaa gct tcg tac ata cac 2515  
 Phe His Ala Tyr Phe Pro Phe Ile Lys Val Glu Ala Ser Tyr Ile His  
 790 795 800 805

caa gat agc ttc aaa gaa cgt aat act acc ttg gta cga tct ttc gat 2563  
 Gln Asp Ser Phe Lys Glu Arg Asn Thr Thr Leu Val Arg Ser Phe Asp  
 810 815 820

agc ggt gat tta att aac gtc tct gtg cct att gga att acc ttc gag 2611  
 Ser Gly Asp Leu Ile Asn Val Ser Val Pro Ile Gly Ile Thr Phe Glu  
 825 830 835

aga ttc tcg aga aac gag cgt gcg tct tac gaa gct act gtc atc tac 2659  
 Arg Phe Ser Arg Asn Glu Arg Ala Ser Tyr Glu Ala Thr Val Ile Tyr  
 840 845 850

gtt gcc gat gtc tat cgt aag aat cct gac tgc acg aca gct ctc cta 2707  
 Val Ala Asp Val Tyr Arg Lys Asn Pro Asp Cys Thr Thr Ala Leu Leu  
 855 860 865

atc aac aat acc tcg tgg aaa act aca gga acg aat ctc tca aga caa 2755  
 Ile Asn Asn Thr Ser Trp Lys Thr Thr Gly Thr Asn Leu Ser Arg Gln  
 870 875 880 885

gct ggt atc gga aga gca ggg atc ttt tat gcc ttc tct cca aat ctt 2803  
 Ala Gly Ile Gly Arg Ala Gly Ile Phe Tyr Ala Phe Ser Pro Asn Leu  
 890 895 900

gag gtc aca agt aac cta tct atg gaa att cgt gga tct tca cgc agc 2851  
 Glu Val Thr Ser Asn Leu Ser Met Glu Ile Arg Gly Ser Ser Arg Ser  
 905 910 915

tac aat gca gat ctt gga ggt aag ttc cag ttc taaaagcgtt cctgatccct 2904  
 Tyr Asn Ala Asp Leu Gly Gly Lys Phe Gln Phe  
 920 925

tagaaattct aagagatcct gagtgtatct agggacttct caaaga 2950

&lt;210&gt;6

&lt;211&gt;2784

&lt;212&gt;DNA

&lt;213&gt;Chlamydia pneumoniae

&lt;220&gt;

&lt;221&gt;CDS

&lt;222&gt;(1)..(2784)

&lt;400&gt;6

|             |            |             |            |             |            |      |
|-------------|------------|-------------|------------|-------------|------------|------|
| atgaaatcct  | ctcttcattg | gtttttaatc  | tcgtcatctt | tagcacttcc  | cttgctacta | 60   |
| aattttctctg | cgtttgctgc | tggtgttgaa  | atcaatctag | gacctacca   | tagcttctct | 120  |
| ggaccaggaa  | cctacactcc | tcagcccaa   | acaacaaatg | cagatggaac  | tatctataat | 180  |
| ctaacagggg  | atgtctcaat | caccaatgca  | ggatctccga | cagctctaac  | cgcttctctg | 240  |
| tttaaagaaa  | ctactgggaa | tctttcttct  | caaggccacg | gtaccaatt   | tctcctacaa | 300  |
| aattctgatg  | cgggagcgaa | ctgtaccttt  | accaatacag | ctgcaaataa  | gcttctctcc | 360  |
| ttttcaggat  | tctcctattt | gtcactaata  | caaaccacga | atgctaccac  | aggaacagga | 420  |
| gccatcaagt  | ccacaggagc | ttgttctatt  | cagtcgaact | atagttgcta  | ctttggccaa | 480  |
| aacttttcta  | atgacaatgg | aggcgccctc  | caaggcagct | ctatcagctt  | atcgctaaac | 540  |
| cccaaccta   | cgtttgccaa | aaacaaagca  | acgcaaaaag | ggggtgccct  | ctattccacg | 600  |
| ggaggggatta | caattaacaa | tacgttaaac  | tcagcatcat | tttctgaaaa  | taccgcggcg | 660  |
| aacaatggcg  | gagccattta | cacggaagct  | agcagtttta | ttagcagcaa  | caaagcaatt | 720  |
| agctttataa  | acaatagctg | gaccgcaacc  | tcagctacag | ggggagccat  | ttactgtagt | 780  |
| agtacatcag  | ccccaaacc  | agtcttaact  | ctatcagaca | acggggaact  | gaactttata | 840  |
| ggaaatacag  | caattactag | tggtggggcg  | atttatactg | acaatctagt  | tctttcttct | 900  |
| ggaggaccta  | cgctttttta | aaacaactct  | ggctatgata | ctgcagctcc  | cttaggagga | 960  |
| gcaattgcga  | ttgctgactc | tggtatctttg | agtctttcgg | ctcttggtgg  | agacatcact | 1020 |
| tttgaaggaa  | acacagtagt | caaaggagct  | tcttcgagtc | agaccactac  | cagaaattct | 1080 |
| attaacatcg  | gaaacaccaa | tgctaagatt  | gtacagctgc | gagcctctca  | aggcaatact | 1140 |
| atctacttct  | atgatcctat | aacaactagc  | atcactgcag | ctctctcaga  | tgctctaaac | 1200 |
| ttaaatggtc  | ctgaccttgc | agggaatcct  | gcatatcaag | gaaccatcgt  | attttctgga | 1260 |
| gagaagctct  | cggaagcaga | agctgcagaa  | gctgataatc | tcaaactctac | aattcagcaa | 1320 |
| cctctaactc  | ttgcggggag | gcaactctct  | cttaaatcag | gagtcactct  | agttgctaag | 1380 |
| tccttttcgc  | aatctccggg | ctctaccctc  | ctcatggatg | cagggaccac  | attagaaacc | 1440 |
| gctgatggga  | tcactatcaa | taatcttgtt  | ctcaatgtag | attccttaaa  | agagaccaag | 1500 |
| aagggcacgc  | taaaagcaac | acaagcaagt  | cagacagtc  | ctttatctgg  | atcgctctct | 1560 |
| cttgtagatc  | cttctggaaa | tgtctacgaa  | gatgtctctt | ggaataacc   | tcaagtcttt | 1620 |
| tcttgctca   | ctcttactgc | tgacgacccc  | gcgaatatc  | acatcacaga  | cttagctgct | 1680 |
| gatcccctag  | aaaaaatcc  | tatccattgg  | ggataccaag | ggaattgggc  | attatcttgg | 1740 |
| caagaggata  | ctgcgactaa | atccaaagca  | gcgactctta | cctggacaaa  | aacaggatac | 1800 |
| aatccgaatc  | ctgagcgctg | tggaacctta  | gttgctaaca | cgctatgggg  | atcctttgtt | 1860 |
| gatgtgcgct  | ccatacaaca | gcttgtagcc  | actaaagtac | gccaatctca  | agaaactcgc | 1920 |
| ggcatctggg  | gtgaagggat | ctcgaaactc  | ttccataaag | atagcacgaa  | gataaataaa | 1980 |
| ggttttcgcc  | acataagtgc | aggttatgtt  | gtaggagcga | ctacaacatt  | agcttctgat | 2040 |
| aatcttatca  | ctgcagcctt | ctgccaatta  | ttcgggaaag | atagagatca  | ctttataaat | 2100 |
| aaaaatagag  | cttctgccta | tgcaagctct  | ctccatctcc | agcatctagc  | gaccttgtct | 2160 |
| tctccaagct  | tgttacgcta | ccttcctgga  | tctgaaagtg | agcagcctgt  | cctctttgat | 2220 |
| gctcagatca  | gctatatcta | tagtaaaaat  | actatgaaaa | cctattacac  | ccaagcacca | 2280 |
| aagggagaga  | gctcgtggta | taatgacggt  | tgcgctctgg | aacttgcgag  | ctccctacca | 2340 |
| cacactgctt  | taagccatga | gggtctcttc  | cacgcgtatt | ttcctttcat  | caaagtagaa | 2400 |
| gcttcgtaca  | tacaccaaga | tagcttcaaa  | gaacgtaata | ctaccttggg  | acgatcttct | 2460 |
| gatagcgggt  | atttaattaa | cgtctctgtg  | cctattggaa | ttaccttcga  | gagattctcg | 2520 |
| agaaacgagc  | gtgcgtctta | cgaagctact  | gtcatctacg | ttgccgatgt  | ctatcgtaag | 2580 |
| aatcctgact  | gcacgacagc | tctcctaata  | aacaatacct | cgtggaaaac  | tacaggaacg | 2640 |
| aatctctcaa  | gacaagctgg | tatcggaaga  | gcagggatct | tttatgcctt  | ctctccaaat | 2700 |
| cttgagggtca | caagtaacct | atctatggaa  | attcgtggat | cttcacgcag  | ctacaatgca | 2760 |
| gatcttggag  | gtaagttcca | gttc        |            |             |            | 2784 |

<210> 7  
 <211> 3000  
 <212> DNA  
 <213> Chlamydia pneumoniae  
 <220>  
 <221> CDS  
 <222> (101)..(2890)  
 <400> 7

```

gtacgaagtt cttcacgaaa ttataataca aacctaggct ctaagttttg tttctagatt 60
atcgaaaacg tgттааттаа ttgaacccaa gcatctttct atg aaa ata ccc ttg 115
                               Met Lys Ile Pro Leu
                               1 5

cac aaa ctc ctg atc tct tcg act ctt gtc act ccc att cta ttg agc 163
His Lys Leu Leu Ile Ser Ser Thr Leu Val Thr Pro Ile Leu Leu Ser
                               10 15 20

att gca act tac gga gca gat gct tct tta tcc cct aca gat agc ttt 211
Ile Ala Thr Tyr Gly Ala Asp Ala Ser Leu Ser Pro Thr Asp Ser Phe
                               25 30 35

gat gga gcg ggc ggc tct aca ttt act cca aaa tct aca gca gat gcc 259
Asp Gly Ala Gly Gly Ser Thr Phe Thr Pro Lys Ser Thr Ala Asp Ala
                               40 45 50

aat gga acg aac tat gtc tta tca gga aat gtc tat ata aac gat gct 307
Asn Gly Thr Asn Tyr Val Leu Ser Gly Asn Val Tyr Ile Asn Asp Ala
                               55 60 65

ggg aaa ggc aca gca tta aca ggc tgc tgc ttt aca gaa act acg ggt 355
Gly Lys Gly Thr Ala Leu Thr Gly Cys Cys Phe Thr Glu Thr Thr Gly
                               70 75 80 85

gat ctg aca ttt act gga aag gga tac tca ttt tca ttc aac acg gta 403
Asp Leu Thr Phe Thr Gly Lys Gly Tyr Ser Phe Ser Phe Asn Thr Val
                               90 95 100

gat gcg ggt tcg aat gca gga gct gcg gca agc aca act gct gat aaa 451
Asp Ala Gly Ser Asn Ala Gly Ala Ala Ala Ser Thr Thr Ala Asp Lys
                               105 110 115

gcc cta atc ttc aca gga ttt tct aac ctt tcc ttc att gca gct cct 499
Ala Leu Ile Phe Thr Gly Phe Ser Asn Leu Ser Phe Ile Ala Ala Pro
                               120 125 130

gga act aca gtt gct tca gga aaa agt act tta agt tct gca gga gcc 547
Gly Thr Thr Val Ala Ser Gly Lys Ser Thr Leu Ser Ser Ala Gly Ala
                               135 140 145

tta aat ctt acc gat aat gga acg att ctc ttt agc caa aac gtc tcc 595
Leu Asn Leu Thr Asp Asn Gly Thr Ile Leu Phe Ser Gln Asn Val Ser
                               150 155 160 165

```

|   |      |
|---|------|
| aat gaa gct aat aac aat ggc gga gcg atc acc aca aaa act ctt tct | 643  |
| Asn Glu Ala Asn Asn Gly Gly Ala Ile Thr Thr Lys Thr Leu Ser     |      |
| 170 175 180   |      |
| att tct ggg aat acc tct tct ata acc ttc act agt aat agc gca aaa | 691  |
| Ile Ser Gly Asn Thr Ser Ser Ile Thr Phe Thr Ser Asn Ser Ala Lys |      |
| 185 190 195   |      |
| aaa tta ggt gga gcg atc tat agc tct gcg gct gca agt att tca gga | 739  |
| Lys Leu Gly Gly Ala Ile Tyr Ser Ser Ala Ala Ala Ser Ile Ser Gly |      |
| 200 205 210   |      |
| aac acc ggc cag tta gtc ttt atg aat aat aaa gga gaa act ggg ggt | 787  |
| Asn Thr Gly Gln Leu Val Phe Met Asn Asn Lys Gly Glu Thr Gly Gly |      |
| 215 220 225   |      |
| ggg gct ctg ggc ttt gaa gcc agc tcc tcg att act caa aat agc tcc | 835  |
| Gly Ala Leu Gly Phe Glu Ala Ser Ser Ser Ile Thr Gln Asn Ser Ser |      |
| 230 235 240 245   |      |
| ctt ttc ttc tct gga aac act gca aca gat gct gca ggc aag ggc ggg | 883  |
| Leu Phe Phe Ser Gly Asn Thr Ala Thr Asp Ala Ala Gly Lys Gly Gly |      |
| 250 255 260   |      |
| gcc att tat tgt gaa aaa aca gga gag act cct act ctt act atc tct | 931  |
| Ala Ile Tyr Cys Glu Lys Thr Gly Glu Thr Pro Thr Leu Thr Ile Ser |      |
| 265 270 275   |      |
| gga aat aaa agt ctg acc ttc gcc gag aac tct tca gta act caa ggc | 979  |
| Gly Asn Lys Ser Leu Thr Phe Ala Glu Asn Ser Ser Val Thr Gln Gly |      |
| 280 285 290   |      |
| gga gca atc tgt gcc cat ggt cta gat ctt tcc gct gct ggc cct acc | 1027 |
| Gly Ala Ile Cys Ala His Gly Leu Asp Leu Ser Ala Ala Gly Pro Thr |      |
| 295 300 305   |      |
| cta ttt tca aat aat aga tgc ggg aac aca gct gca ggc aag ggc ggc | 1075 |
| Leu Phe Ser Asn Asn Arg Cys Gly Asn Thr Ala Ala Gly Lys Gly Gly |      |
| 310 315 320 325   |      |
| gct att gca att gcc gac tct gga tct tta agt ctc tct gca aat caa | 1123 |
| Ala Ile Ala Ile Ala Asp Ser Gly Ser Leu Ser Leu Ser Ala Asn Gln |      |
| 330 335 340   |      |
| gga gac atc acg ttc ctt ggc aac act cta acc tca acc tcc gcg cca | 1171 |
| Gly Asp Ile Thr Phe Leu Gly Asn Thr Leu Thr Ser Thr Ser Ala Pro |      |
| 345 350 355   |      |
| aca tcg aca cgg aat gct atc tac ctg gga tcg tca gca aaa att acg | 1219 |
| Thr Ser Thr Arg Asn Ala Ile Tyr Leu Gly Ser Ser Ala Lys Ile Thr |      |
| 360 365 370   |      |
| aac tta agg gca gcc caa ggc caa tct atc tat ttc tat gat ccg att | 1267 |
| Asn Leu Arg Ala Ala Gln Gly Gln Ser Ile Tyr Phe Tyr Asp Pro Ile |      |
| 375 380 385   |      |

|   |      |
|---|------|
| gca tct aac acc aca gga gct tca gac gtt ctg acc atc aac caa ccg | 1315 |
| Ala Ser Asn Thr Thr Gly Ala Ser Asp Val Leu Thr Ile Asn Gln Pro |      |
| 390 395 400 405   |      |
| gat agc aac tcg cct tta gat tat tca gga acg att gta ttt tct ggg | 1363 |
| Asp Ser Asn Ser Pro Leu Asp Tyr Ser Gly Thr Ile Val Phe Ser Gly |      |
| 410 415 420   |      |
| gaa aag ctc tct gca gat gaa gcg aaa gct gct gat aac ttc aca tct | 1411 |
| Glu Lys Leu Ser Ala Asp Glu Ala Lys Ala Ala Asp Asn Phe Thr Ser |      |
| 425 430 435   |      |
| ata tta aag caa cca ttg gct cta gcc tct gga acc tta gca ctc aaa | 1459 |
| Ile Leu Lys Gln Pro Leu Ala Leu Ala Ser Gly Thr Leu Ala Leu Lys |      |
| 440 445 450   |      |
| gga aat gtc gag tta gat gtc aat ggt ttc aca cag act gaa ggc tct | 1507 |
| Gly Asn Val Glu Leu Asp Val Asn Gly Phe Thr Gln Thr Glu Gly Ser |      |
| 455 460 465   |      |
| aca ctc ctc atg caa cca gga aca aag ctc aaa gca gat act gaa gct | 1555 |
| Thr Leu Leu Met Gln Pro Gly Thr Lys Leu Lys Ala Asp Thr Glu Ala |      |
| 470 475 480 485   |      |
| atc agt ctt acc aaa ctt gtc gtt gat ctt tct gcc tta gag gga aat | 1603 |
| Ile Ser Leu Thr Lys Leu Val Val Asp Leu Ser Ala Leu Glu Gly Asn |      |
| 490 495 500   |      |
| aag agt gtg tcc att gaa aca gca gga gcc aac aaa act ata act cta | 1651 |
| Lys Ser Val Ser Ile Glu Thr Ala Gly Ala Asn Lys Thr Ile Thr Leu |      |
| 505 510 515   |      |
| acc tct cct ctt gtt ttc caa gat agt agc ggc aat ttt tat gaa agc | 1699 |
| Thr Ser Pro Leu Val Phe Gln Asp Ser Ser Gly Asn Phe Tyr Glu Ser |      |
| 520 525 530   |      |
| cat acg ata aac caa gcc ttc acg cag cct ttg gtg gta ttc act gct | 1747 |
| His Thr Ile Asn Gln Ala Phe Thr Gln Pro Leu Val Val Phe Thr Ala |      |
| 535 540 545   |      |
| gct act gct gct agc gat att tat atc gat gcg ctt ctc act tct cca | 1795 |
| Ala Thr Ala Ala Ser Asp Ile Tyr Ile Asp Ala Leu Leu Thr Ser Pro |      |
| 550 555 560 565   |      |
| gta caa act cca gaa cct cat tac ggg tat cag gga cat tgg gaa gcc | 1843 |
| Val Gln Thr Pro Glu Pro His Tyr Gly Tyr Gln Gly His Trp Glu Ala |      |
| 570 575 580   |      |
| act tgg gca gac aca tca act gca aaa tca gga act atg act tgg gta | 1891 |
| Thr Trp Ala Asp Thr Ser Thr Ala Lys Ser Gly Thr Met Thr Trp Val |      |
| 585 590 595   |      |
| act acg ggc tac aac cct aat cct gag cgt aga gct tcc gta gtt ccc | 1939 |
| Thr Thr Gly Tyr Asn Pro Asn Pro Glu Arg Arg Ala Ser Val Val Pro |      |
| 600 605 610   |      |



|   |      |
|---|------|
| gat tca tta tgg gca tcc ttt act gac att cgc act cta cag cag atc<br>Asp Ser Leu Trp Ala Ser Phe Thr Asp Ile Arg Thr Leu Gln Gln Ile<br>615 620 625     | 1987 |
| atg aca tct caa gcg aat agt atc tat cag caa cga gga ctc tgg gca<br>Met Thr Ser Gln Ala Asn Ser Ile Tyr Gln Gln Arg Gly Leu Trp Ala<br>630 635 640 645 | 2035 |
| tca gga act gcg aat ttc ttc cat aag gat aaa tca gga act aac caa<br>Ser Gly Thr Ala Asn Phe Phe His Lys Asp Lys Ser Gly Thr Asn Gln<br>650 655 660     | 2083 |
| gca ttc cga cat aaa agc tac ggc tat att gtt gga gga agt gct gaa<br>Ala Phe Arg His Lys Ser Tyr Gly Tyr Ile Val Gly Gly Ser Ala Glu<br>665 670 675     | 2131 |
| gat ttt tct gaa aat atc ttc agt gta gct ttc tgc cag ctc ttc ggt<br>Asp Phe Ser Glu Asn Ile Phe Ser Val Ala Phe Cys Gln Leu Phe Gly<br>680 685 690     | 2179 |
| aaa gat aaa gac ctg ttt ata gtt gaa aat acc tct cat aac tat tta<br>Lys Asp Lys Asp Leu Phe Ile Val Glu Asn Thr Ser His Asn Tyr Leu<br>695 700 705     | 2227 |
| gcg tcg cta tac ctg caa cat cga gca ttc cta gga gga ctt ccc atg<br>Ala Ser Leu Tyr Leu Gln His Arg Ala Phe Leu Gly Gly Leu Pro Met<br>710 715 720 725 | 2275 |
| ccc tca ttt gga agt atc acc gac atg ctg aaa gat att cct ctc att<br>Pro Ser Phe Gly Ser Ile Thr Asp Met Leu Lys Asp Ile Pro Leu Ile<br>730 735 740     | 2323 |
| ttg aat gcc cag cta agc tac agc tac act aaa aat gat atg gat act<br>Leu Asn Ala Gln Leu Ser Tyr Ser Tyr Thr Lys Asn Asp Met Asp Thr<br>745 750 755     | 2371 |
| cgc tat act tcc tat cct gaa gct caa ggc tct tgg acc aat aac tct<br>Arg Tyr Thr Ser Tyr Pro Glu Ala Gln Gly Ser Trp Thr Asn Asn Ser<br>760 765 770     | 2419 |
| ggg gct cta gag ctc gga gga tct ctg gct cta tat ctc cct aaa gaa<br>Gly Ala Leu Glu Leu Gly Gly Ser Leu Ala Leu Tyr Leu Pro Lys Glu<br>775 780 785     | 2467 |
| gca ccg ttc ttc cag gga tat ttc ccc ttc tta aag ttc cag gca gtc<br>Ala Pro Phe Phe Gln Gly Tyr Phe Pro Phe Leu Lys Phe Gln Ala Val<br>790 795 800 805 | 2515 |
| tac agc cgc caa caa aac ttt aaa gag agt ggc gct gaa gcc cgt gct<br>Tyr Ser Arg Gln Gln Asn Phe Lys Glu Ser Gly Ala Glu Ala Arg Ala<br>810 815 820     | 2563 |
| ttt gat gat gga gac cta gtg aac tgc tct atc cct gtc ggc att cgg<br>Phe Asp Asp Gly Asp Leu Val Asn Cys Ser Ile Pro Val Gly Ile Arg<br>825 830 835     | 2611 |

tta gaa aaa atc tcc gaa gat gaa aaa aat aat ttc gag att tct cta 2659  
 Leu Glu Lys Ile Ser Glu Asp Glu Lys Asn Asn Phe Glu Ile Ser Leu  
 840 845 850  
 gcc tac att ggt gat gtg tat cgt aaa aat ccc cgt tcg cgt act tct 2707  
 Ala Tyr Ile Gly Asp Val Tyr Arg Lys Asn Pro Arg Ser Arg Thr Ser  
 855 860 865  
 cta atg gtc agt gga gcc tct tgg act tcg cta tgt aaa aac ctc gca 2755  
 Leu Met Val Ser Gly Ala Ser Trp Thr Ser Leu Cys Lys Asn Leu Ala  
 870 875 880 885  
 cga caa gcc ttc tta gca agt gct gga agc cat ctg act ctc tcc cct 2803  
 Arg Gln Ala Phe Leu Ala Ser Ala Gly Ser His Leu Thr Leu Ser Pro  
 890 895 900  
 cat gta gaa ctc tct ggg gaa gct gct tat gag ctt cgt ggc tca gca 2851  
 His Val Glu Leu Ser Gly Glu Ala Ala Tyr Glu Leu Arg Gly Ser Ala  
 905 910 915  
 cac atc tac aat gta gat tgt ggg cta aga tac tca ttc tagttcctac 2900  
 His Ile Tyr Asn Val Asp Cys Gly Leu Arg Tyr Ser Phe  
 920 925 930  
 tttcctccct aaacttttag ggaggaattc ttataaaaac cctgtagatt cttaacttac 2960  
 tagtctctcc tttcctcttg ctttctttaa tttattgcag 3000

&lt;210&gt;8

&lt;211&gt;2790

&lt;212&gt;DNA

&lt;213&gt;Chlamydia pneumoniae

&lt;220&gt;

&lt;221&gt;CDS

&lt;222&gt; (1) .. (2790)

&lt;400&gt;8

atgaaaatac ccttgacaaa actcctgac tcttcgactc ttgtcactcc cattctattg 60  
 agcattgcaa cttacggagc agatgcttct ttatccccta cagatagctt tgatggagcg 120  
 ggcggctcta catttactcc aaaatctaca gcagatgcc aatggaacgaa ctatgtctta 180  
 tcaggaaatg tctatataaa cgatgctggg aaaggcacag cattaacagg ctgctgcttt 240  
 acagaaacta cgggtgatct gacatttact ggaaagggat actcattttc attcaacacg 300  
 gtagatgcgg gttcgaatgc aggagctgcg gcaagcacia ctgctgataa agccctaate 360  
 ttcacaggat tttctaacct ttccttcatt gcagctcctg gaactacagt tgcttcagga 420  
 aaaagtactt taagttctgc aggagcctta aatcttaccg ataatggaac gattctcttt 480  
 agccaaaacg tctccaatga agctaataac aatggcggag cgatcaccac aaaaactctt 540  
 tctattttctg ggaataacct ttctataacc ttcactagta atagcgcaaa aaaattaggt 600  
 ggagcgatct atagctctgc ggctgcaagt atttcaggaa acaccggcca gttagtcttt 660  
 atgaataata aaggagaaac tgggggtggg gctctgggct ttgaagccag ctccctcgatt 720  
 actcaaaata gctccctttt cttctctgga aacactgcaa cagatgctgc aggcaagggc 780  
 ggggccattt attgtgaaaa aacaggagag actcctactc ttactatctc tggaaataaa 840  
 agtctgacct tcgccgagaa ctcttcagta actcaaggcg gagcaatctg tgcccatggt 900  
 ctatgctttt ccgctgctgg ccctacccta ttttcaaata atagatgcgg gaacacagct 960  
 gcaggcaagg gcggcgctat tgcaattgcc gactctggat ctttaagtct ctctgcaaat 1020

```

caaggagaca tcacgttcct tggcaacact ctaacctcaa cctccgcgcc aacatcgaca 1080
cggaatgcta tctacctggg atcgtcagca aaaattacga acttaagggc agcccaaggc 1140
caatctatct atttctatga tccgattgca tctaaccacca caggagcttc agacgttctg 1200
accatcaacc aaccggatag caactcgctt ttagattatt caggaacgat tgtattttct 1260
ggggaaaagc tctctgcaga tgaagcgaaa gctgctgata acttcacatc tatattaaag 1320
caaccattgg ctctagcctc tggaaacctt gcaactcaaag gaaatgtcga gttagatgtc 1380
aatggtttca cacagactga aggtcttaca ctctcatgac aaccaggaac aaagctcaaa 1440
gcagatactg aagctatcag tcttaccaaa cttgtcgttg atctttctgc cttagaggga 1500
aataagagtg tgtccattga aacagcagga gccacaacaa ctataactct aacctctcct 1560
cttgttttcc aagatagtag cggcaatttt tatgaaagcc atacgataaa ccaagccttc 1620
acgcagcctt tgggtgtatt actgtgcta ctgctgctag cgatatttat atcgatgcgc 1680
ttctcacttc tccagtacaa actccagaac ctcatcagg gtatcaggga cattgggaag 1740
ccacttgggc agacacatca actgcaaaat caggaaactat gacttgggta actacgggct 1800
acaaccctaa tcctgagcgt agagcttccg tagttcccga ttcattatgg gcatccttta 1860
ctgacattcg cactctacag cagatcatga catctcaagc gaatagtatc tatcagcaac 1920
gaggactctg ggcacagga actgcgaatt tcttccataa ggataaatca ggaactaacc 1980
aagcattccg acataaaaag tacggctata ttgttggagg aagtgtgaa gatttttctg 2040
aaaatatctt cagtgtatgt ttctgccagc tcttcggtaa agataaagac ctgtttatag 2100
ttgaaaatac ctctcataac tatttagcgt cgctataact gcaacatcga gcattcctag 2160
gaggacttcc catgccctca tttggaagta tcaccgacat gctgaaagat attcctctca 2220
ttttgaatgc ccagctaagc tacagctaca ctaaaaatga tatggatact cgctatactt 2280
cctatcctga agctcaaggc tcttggacca ataaactctg ggctctagag ctcgagggat 2340
ctctggctct atatctccct aaagaagcac cgttcttcca gggatatttc cccttcttaa 2400
agtctcagc agtctacagc cgccaacaaa actttaaaga gagtggcgct gaagcccggt 2460
cttttgatga tggagaccta gtgaactgct ctatccctgt cggcattcgg ttagaaaaaa 2520
tctccgaaga tgaaaaaaat aatttcgaga tttctctagc ctacattggg gatgtgtatc 2580
gtaaaaatcc ccgttcgcgt acttctctaa tggctcagc ggagcctctt ggacttcgct 2640
tagtaaaaac ctgcacgac aagccttctt agcaagtgtc ggaagccatc tgactctctc 2700
cctcatgtag aactctctgg ggaagctgct tatgagcttc gtggctcagc acacatctac 2760
aatgtagatt gtgggctaag atactcattc 2790

```

```

<210> 9
<211> 1100
<212> DNA
<213> Chlamydia pneumoniae
<220>
<221> CDS
<222> (101)..(979)
<400> 9

```

```

ataaagtttt ttatatgaac aaaactttga cttcattggc agacacttcg accttaacag 60

accaattttg ttgtcatccc tataaaaatc aggaattttc atg ctc tcc tca cta 115
                                     Met Leu Ser Ser Leu
                                     1                               5

atc cgt gat tca ttt ccc ctt ctt att tta ctt ccc aca ttc cta gcg 163
Ile Arg Asp Ser Phe Pro Leu Leu Ile Leu Leu Pro Thr Phe Leu Ala
                10                        15                        20

gca tta gga gcc tcc gta gct ggc ggc gtt atg gga acc tat atc gtt 211
Ala Leu Gly Ala Ser Val Ala Gly Gly Val Met Gly Thr Tyr Ile Val
                25                        30                        35

```

|   |     |
|---|-----|
| gta aaa cgt att gtt tca att agt gga agt ata tct cat gca att cta | 259 |
| Val Lys Arg Ile Val Ser Ile Ser Gly Ser Ile Ser His Ala Ile Leu |     |
| 40 45 50  |     |
| gga gga att ggc ctc acc cta tgg ata caa tat aag ctt cat ctc tct | 307 |
| Gly Gly Ile Gly Leu Thr Leu Trp Ile Gln Tyr Lys Leu His Leu Ser |     |
| 55 60 65  |     |
| ttt ttc cct atg tat gga gct att gta gga gct att ttt cta gct ctt | 355 |
| Phe Phe Pro Met Tyr Gly Ala Ile Val Gly Ala Ile Phe Leu Ala Leu |     |
| 70 75 80 85   |     |
| tgc atc ggc aag atc cac ctg aaa tac caa gaa agg gaa gac tct ttg | 403 |
| Cys Ile Gly Lys Ile His Leu Lys Tyr Gln Glu Arg Glu Asp Ser Leu |     |
| 90 95 100   |     |
| att gcg atg att tgg tct gtg ggc atg gca att gga att ata ttc att | 451 |
| Ile Ala Met Ile Trp Ser Val Gly Met Ala Ile Gly Ile Ile Phe Ile |     |
| 105 110 115   |     |
| tcc agg ctt ccc acc ttt aat gga gag ctc atc aat ttt cta ttt ggg | 499 |
| Ser Arg Leu Pro Thr Phe Asn Gly Glu Leu Ile Asn Phe Leu Phe Gly |     |
| 120 125 130   |     |
| aac att ctc tgg gtc acc cct tca gac ctc tat agc tta gga atc ttt | 547 |
| Asn Ile Leu Trp Val Thr Pro Ser Asp Leu Tyr Ser Leu Gly Ile Phe |     |
| 135 140 145   |     |
| gat ctt ctt gtt tta gga att gtg gtc ctt tgc cac acc cgg ttc ctt | 595 |
| Asp Leu Leu Val Leu Gly Ile Val Val Leu Cys His Thr Arg Phe Leu |     |
| 150 155 160 165   |     |
| gct ctt tgc ttt gat gag agg tac acg gct tta aac cat tgt tct gta | 643 |
| Ala Leu Cys Phe Asp Glu Arg Tyr Thr Ala Leu Asn His Cys Ser Val |     |
| 170 175 180   |     |
| cag ctg tgg tat ttc cta ctt ctt gtt ctg aca gca atc acg att gtg | 691 |
| Gln Leu Trp Tyr Phe Leu Leu Leu Val Leu Thr Ala Ile Thr Ile Val |     |
| 185 190 195   |     |
| atg ttg att tat gtg atg gga acg att ctg atg ctt agc atg ctc gtc | 739 |
| Met Leu Ile Tyr Val Met Gly Thr Ile Leu Met Leu Ser Met Leu Val |     |
| 200 205 210   |     |
| tta cct gtt gct ata gcg tgt aga ttt tcg tac aag atg aca cga att | 787 |
| Leu Pro Val Ala Ile Ala Cys Arg Phe Ser Tyr Lys Met Thr Arg Ile |     |
| 215 220 225   |     |
| atg ttc atc tcg gtc ctc ttg aat atc tta tgt tct ttt tct gga att | 835 |
| Met Phe Ile Ser Val Leu Leu Asn Ile Leu Cys Ser Phe Ser Gly Ile |     |
| 230 235 240 245   |     |
| tgc atc gcc tac tgt tta gat ttc cca gta ggt cct acg ata tca ttg | 883 |
| Cys Ile Ala Tyr Cys Leu Asp Phe Pro Val Gly Pro Thr Ile Ser Leu |     |
| 250 255 260   |     |

ctg atg ggg tta ggt tat aca gcg agt ctt tgt gtg aag aag cgg tac 931  
 Leu Met Gly Leu Gly Tyr Thr Ala Ser Leu Cys Val Lys Lys Arg Tyr  
           265                                  270                                  275  
  
 aat ccg tcg acg cct tct cct gta agt cct gaa atc aat aca aat gta 979  
 Asn Pro Ser Thr Pro Ser Pro Val Ser Pro Glu Ile Asn Thr Asn Val  
           280                                  285                                  290  
  
 tagctaggga agcgcttttg gaagctttgg aggcattctt cctgttcgtc aggaagaaga 1039  
 tcatcaattt tatttaaagc taccagcata tctttctttt caaaatctgg ctgatgagag 1099  
 t 1100

<210> 10  
 <211> 880  
 <212> DNA  
 <213> Chlamydiapneumoniae  
 <220>  
 <221> CDS  
 <222> (1)..(880)  
 <400> 10

atgctctcct cactaatccg tgattcattt ccccttctta ttttacttcc cacattccta 60  
 gcggcattag gagcctccgt agctggcggc gttatgggaa cctatatcgt tgtaaaacgt 120  
 attgtttcaa ttagtggaag tatactctcat gcaattctag gaggaattgg cctcacccta 180  
 tggatacaat ataagcttca tctctctttt ttccctatgt atggagctat tgtaggagct 240  
 atttttctag ctctttgcat cggcaagatc cacctgaaat accaagaaag ggaagactct 300  
 ttgattgcga tgatttggtc tgtgggcatg gcaattggaa ttatattcat' ttccaggctt 360  
 cccaccttta atggagagct catcaatttt ctatttgga acattctctg ggtcaccctt 420  
 tcagacctct atagcttagg aatctttgat cttcttggtt taggaattgt ggtcctttgc 480  
 cacacccggg tccttgctct ttgctttgat gagaggtaca cggttttaaa ccattgttct 540  
 gtacagctgt ggtatttctt acttcttggt ctgacagcaa tcacgattgt gatgttgatt 600  
 tatgtgatgg gaacgattct gatgcttagc atgctcgtct tacctgttgc tatagcgtgt 660  
 agattttcgt acaagatgac acgaattatg ttcattctcg tcctcttgaa tatcttatgt 720  
 tctttttctg gaatttgcat cgctactgt ttagatttcc cagtaggtcc tacgatatca 780  
 ttgctgatgg ggttaggtta tacagcgagt ctttgtgtag aagaagcggg acaatccgtc 840  
 gacgccttct cctgtaagtc ctgaaatcaa tacaatgta 880

<210> 11  
 <211> 928  
 <212> PRT  
 <213> Chlamydia pneumoniae  
 <400> 11

Met Lys Thr Ser Ile Pro Trp Val Leu Val Ser Ser Val Leu Ala Phe  
   1                                  5                                  10                                  15  
  
 Ser Cys His Leu Gln Ser Leu Ala Asn Glu Glu Leu Leu Ser Pro Asp  
           20                                  25                                  30

Asp Ser Phe Asn Gly Asn Ile Asp Ser Gly Thr Phe Thr Pro Lys Thr  
 35 40 45  
 Ser Ala Thr Thr Tyr Ser Leu Thr Gly Asp Val Phe Phe Tyr Glu Pro  
 50 55 60  
 Gly Lys Gly Thr Pro Leu Ser Asp Ser Cys Phe Lys Gln Thr Thr Asp  
 65 70 75 80  
 Asn Leu Thr Phe Leu Gly Asn Gly His Ser Leu Thr Phe Gly Phe Ile  
 85 90 95  
 Asp Ala Gly Thr His Ala Gly Ala Ala Ala Ser Thr Thr Ala Asn Lys  
 100 105 110  
 Asn Leu Thr Phe Ser Gly Phe Ser Leu Leu Ser Phe Asp Ser Ser Pro  
 115 120 125  
 Ser Thr Thr Val Thr Thr Gly Gln Gly Thr Leu Ser Ser Ala Gly Gly  
 130 135 140  
 Val Asn Leu Glu Asn Ile Arg Lys Leu Val Val Ala Gly Asn Phe Ser  
 145 150 155 160  
 Thr Ala Asp Gly Gly Ala Ile Lys Gly Ala Ser Phe Leu Leu Thr Gly  
 165 170 175  
 Thr Ser Gly Asp Ala Leu Phe Ser Asn Asn Ser Ser Ser Thr Lys Gly  
 180 185 190  
 Gly Ala Ile Ala Thr Thr Ala Gly Ala Arg Ile Ala Asn Asn Thr Gly  
 195 200 205  
 Tyr Val Arg Phe Leu Ser Asn Ile Ala Ser Thr Ser Gly Gly Ala Ile  
 210 215 220  
 Asp Asp Glu Gly Thr Ser Ile Leu Ser Asn Asn Lys Phe Leu Tyr Phe  
 225 230 235 240  
 Glu Gly Asn Ala Ala Lys Thr Thr Gly Gly Ala Ile Cys Asn Thr Lys  
 245 250 255  
 Ala Ser Gly Ser Pro Glu Leu Ile Ile Ser Asn Asn Lys Thr Leu Ile  
 260 265 270  
 Phe Ala Ser Asn Val Ala Glu Thr Ser Gly Gly Ala Ile His Ala Lys  
 275 280 285  
 Lys Leu Ala Leu Ser Ser Gly Gly Phe Thr Glu Phe Leu Arg Asn Asn  
 290 295 300  
 Val Ser Ser Ala Thr Pro Lys Gly Gly Ala Ile Ser Ile Asp Ala Ser  
 305 310 315 320  
 Gly Glu Leu Ser Leu Ser Ala Glu Thr Gly Asn Ile Thr Phe Val Arg  
 325 330 335

Asn Thr Leu Thr Thr Thr Gly Ser Thr Asp Thr Pro Lys Arg Asn Ala  
 340 345 350  
 Ile Asn Ile Gly Ser Asn Gly Lys Phe Thr Glu Leu Arg Ala Ala Lys  
 355 360 365  
 Asn His Thr Ile Phe Phe Tyr Asp Pro Ile Thr Ser Glu Gly Thr Ser  
 370 375 380  
 Ser Asp Val Leu Lys Ile Asn Asn Gly Ser Ala Gly Ala Leu Asn Pro  
 385 390 395 400  
 Tyr Gln Gly Thr Ile Leu Phe Ser Gly Glu Thr Leu Thr Ala Asp Glu  
 405 410 415  
 Leu Lys Val Ala Asp Asn Leu Lys Ser Ser Phe Thr Gln Pro Val Ser  
 420 425 430  
 Leu Ser Gly Gly Lys Leu Leu Leu Gln Lys Gly Val Thr Leu Glu Ser  
 435 440 445  
 Thr Ser Phe Ser Gln Glu Ala Gly Ser Leu Leu Gly Met Asp Ser Gly  
 450 455 460  
 Thr Thr Leu Ser Thr Thr Ala Gly Ser Ile Thr Ile Thr Asn Leu Gly  
 465 470 475 480  
 Ile Asn Val Asp Ser Leu Gly Leu Lys Gln Pro Val Ser Leu Thr Ala  
 485 490 495  
 Lys Gly Ala Ser Asn Lys Val Ile Val Ser Gly Lys Leu Asn Leu Ile  
 500 505 510  
 Asp Ile Glu Gly Asn Ile Tyr Glu Ser His Met Phe Ser His Asp Gln  
 515 520 525  
 Leu Phe Ser Leu Leu Lys Ile Thr Val Asp Ala Asp Val Asp Thr Asn  
 530 535 540  
 Val Asp Ile Ser Ser Leu Ile Pro Val Pro Ala Glu Asp Pro Asn Ser  
 545 550 555 560  
 Glu Tyr Gly Phe Gln Gly Gln Trp Asn Val Asn Trp Thr Thr Asp Thr  
 565 570 575  
 Ala Thr Asn Thr Lys Glu Ala Thr Ala Thr Trp Thr Lys Thr Gly Phe  
 580 585 590  
 Val Pro Ser Pro Glu Arg Lys Ser Ala Leu Val Cys Asn Thr Leu Trp  
 595 600 605  
 Gly Val Phe Thr Asp Ile Arg Ser Leu Gln Gln Leu Val Glu Ile Gly  
 610 615 620  
 Ala Thr Gly Met Glu His Lys Gln Gly Phe Trp Val Ser Ser Met Thr  
 625 630 635 640

Asn Phe Leu His Lys Thr Gly Asp Glu Asn Arg Lys Gly Phe Arg His  
 645 650 655  
 Thr Ser Gly Gly Tyr Val Ile Gly Gly Ser Ala His Thr Pro Lys Asp  
 660 665 670  
 Asp Leu Phe Thr Phe Ala Phe Cys His Leu Phe Ala Arg Asp Lys Asp  
 675 680 685  
 Cys Phe Ile Ala His Asn Asn Ser Arg Thr Tyr Gly Gly Thr Leu Phe  
 690 695 700  
 Phe Lys His Ser His Thr Leu Gln Pro Gln Asn Tyr Leu Arg Leu Gly  
 705 710 715 720  
 Arg Ala Lys Phe Ser Glu Ser Ala Ile Glu Lys Phe Pro Arg Glu Ile  
 725 730 735  
 Pro Leu Ala Leu Asp Val Gln Val Ser Phe Ser His Ser Asp Asn Arg  
 740 745 750  
 Met Glu Thr His Tyr Thr Ser Leu Pro Glu Ser Glu Gly Ser Trp Ser  
 755 760 765  
 Asn Glu Cys Ile Ala Gly Gly Ile Gly Leu Asp Leu Pro Phe Val Leu  
 770 775 780  
 Ser Asn Pro His Pro Leu Phe Lys Thr Phe Ile Pro Gln Met Lys Val  
 785 790 795 800  
 Glu Met Val Tyr Val Ser Gln Asn Ser Phe Phe Glu Ser Ser Ser Asp  
 805 810 815  
 Gly Arg Gly Phe Ser Ile Gly Arg Leu Leu Asn Leu Ser Ile Pro Val  
 820 825 830  
 Gly Ala Lys Phe Val Gln Gly Asp Ile Gly Asp Ser Tyr Thr Tyr Asp  
 835 840 845  
 Leu Ser Gly Phe Phe Val Ser Asp Val Tyr Arg Asn Asn Pro Gln Ser  
 850 855 860  
 Thr Ala Thr Leu Val Met Ser Pro Asp Ser Trp Lys Ile Arg Gly Gly  
 865 870 875 880  
 Asn Leu Ser Arg Gln Ala Phe Leu Leu Arg Gly Ser Asn Asn Tyr Val  
 885 890 895  
 Tyr Asn Ser Asn Cys Glu Leu Phe Gly His Tyr Ala Met Glu Leu Arg  
 900 905 910  
 Gly Ser Ser Arg Asn Tyr Asn Val Asp Val Gly Thr Lys Leu Arg Phe  
 915 920 925



<210> 13  
 <211> 885  
 <212> PRT  
 <213> Chlamydia pneumoniae  
 <400> 13

Gly Thr Tyr Thr Pro Lys Asn Thr Thr Thr Gly Ile Asp Tyr Thr Leu  
 1 5 10 15

Thr Gly Asp Ile Thr Leu Gln Asn Leu Gly Asp Ser Ala Ala Leu Thr  
 20 25 30

Lys Gly Cys Phe Ser Asp Thr Thr Glu Ser Leu Ser Phe Ala Gly Lys  
 35 40 45

Gly Tyr Ser Leu Ser Phe Leu Asn Ile Lys Ser Ser Ala Glu Gly Ala  
 50 55 60

Ala Leu Ser Val Thr Thr Asp Lys Asn Leu Ser Leu Thr Gly Phe Ser  
 65 70 75 80

Ser Leu Thr Phe Leu Ala Ala Pro Ser Ser Val Ile Thr Thr Pro Ser  
 85 90 95

Gly Lys Gly Ala Val Lys Cys Gly Gly Asp Leu Thr Phe Asp Asn Asn  
 100 105 110

Gly Thr Ile Leu Phe Lys Gln Asp Tyr Cys Glu Glu Asn Gly Gly Ala  
 115 120 125

Ile Ser Thr Lys Asn Leu Ser Leu Lys Asn Ser Thr Gly Ser Ile Ser  
 130 135 140

Phe Glu Gly Asn Lys Ser Ser Ala Thr Gly Lys Lys Gly Gly Ala Ile  
 145 150 155 160

Cys Ala Thr Gly Thr Val Asp Ile Thr Asn Asn Thr Ala Pro Thr Leu  
 165 170 175

Phe Ser Asn Asn Ile Ala Glu Ala Ala Gly Gly Ala Ile Asn Ser Thr  
 180 185 190

Gly Asn Cys Thr Ile Thr Gly Asn Thr Ser Leu Val Phe Ser Glu Asn  
 195 200 205

Ser Val Thr Ala Thr Ala Gly Asn Gly Gly Ala Leu Ser Gly Asp Ala  
 210 215 220

Asp Val Thr Ile Ser Gly Asn Gln Ser Val Thr Phe Ser Gly Asn Gln  
 225 230 235 240

Ala Val Ala Asn Gly Gly Ala Ile Tyr Ala Lys Lys Leu Thr Leu Ala  
 245 250 255

Ser Gly Gly Gly Gly Gly Asn Pro Phe Ser Asn Asn Ile Val Gln Gly  
 260 265 270

Thr Thr Ala Gly Asn Gly Gly Ala Ile Ser Ile Leu Ala Ala Gly Glu  
 275 280 285  
 Cys Ser Leu Phe Ser Glu Ala Gly Asp His Tyr Leu Asn Gly Asn Ala  
 290 295 300  
 Ile Val Ala Thr Thr Pro Gln Thr Thr Lys Arg Asn Ser Ile Asp Ile  
 305 310 315 320  
 Gly Ser Thr Gly Lys Asp His Glu Leu Arg Ala Ile Ser Gly His Ser  
 325 330 335  
 Ile Phe Phe Tyr Asp Pro Ile Thr Ala Asn Thr Ala Ala Asp Ser Thr  
 340 345 350  
 Asp Thr Leu Asn Leu Asn Lys Ala Asp Ala Gly Asn Ser Thr Asp Tyr  
 355 360 365  
 Ser Gly Ser Ile Val Phe Ser Gly Glu Lys Leu Ser Glu Asp Glu Ala  
 370 375 380  
 Lys Val Ala Asp Asn Leu Thr Ser Thr Leu Lys Gln Pro Val Thr Leu  
 385 390 395 400  
 Thr Ala Gly Asn Leu Val Leu Lys Arg Gly Val Thr Leu Asp Thr Lys  
 405 410 415  
 Gly Phe Thr Gln Thr Ala Gly Ser Ser Val Ile Met Asp Ala Gly Thr  
 420 425 430  
 Thr Leu Lys Ala Ser Thr Glu Glu Val Thr Leu Thr Gly Leu Ser Ile  
 435 440 445  
 Pro Val Asp Ser Leu Gly Glu Gly Lys Lys Val Val Ile Ala Ala Ser  
 450 455 460  
 Ala Ala Ser Lys Asn Val Ala Leu Ser Gly Pro Ile Leu Leu Leu Asp  
 465 470 475 480  
 Asn Gln Gly Asn Ala Tyr Glu Asn His Asp Leu Gly Lys Thr Gln Asp  
 485 490 495  
 Phe Ser Phe Val Gln Leu Ser Ala Leu Gly Thr Ala Thr Thr Thr Asp  
 500 505 510  
 Val Pro Ala Val Pro Thr Val Ala Thr Pro Thr His Tyr Gly Tyr Gln  
 515 520 525  
 Gly Thr Trp Gly Met Thr Trp Val Asp Asp Thr Ala Ser Thr Pro Lys  
 530 535 540  
 Thr Lys Thr Ala Thr Leu Ala Trp Thr Asn Thr Gly Tyr Leu Pro Asn  
 545 550 555 560  
 Pro Glu Arg Gln Gly Pro Leu Val Pro Asn Ser Leu Trp Gly Ser Phe  
 565 570 575

Ser Asp Ile Gln Ala Ile Gln Gly Val Ile Glu Arg Ser Ala Leu Thr  
 580 585 590  
 Leu Cys Ser Asp Arg Gly Phe Trp Ala Ala Gly Val Ala Asn Phe Leu  
 595 600 605  
 Asp Lys Asp Lys Lys Gly Glu Lys Arg Lys Tyr Arg His Lys Ser Gly  
 610 615 620  
 Gly Tyr Ala Ile Gly Gly Ala Ala Gln Thr Cys Ser Glu Asn Leu Ile  
 625 630 635 640  
 Ser Phe Ala Phe Cys Gln Leu Phe Gly Ser Asp Lys Asp Phe Leu Val  
 645 650 655  
 Ala Lys Asn His Thr Asp Thr Tyr Ala Gly Ala Phe Tyr Ile Gln His  
 660 665 670  
 Ile Thr Glu Cys Ser Gly Phe Ile Gly Cys Leu Leu Asp Lys Leu Pro  
 675 680 685  
 Gly Ser Trp Ser His Lys Pro Leu Val Leu Glu Gly Gln Leu Ala Tyr  
 690 695 700  
 Ser His Val Ser Asn Asp Leu Lys Thr Lys Tyr Thr Ala Tyr Pro Glu  
 705 710 715 720  
 Val Lys Gly Ser Trp Gly Asn Asn Ala Phe Asn Met Met Leu Gly Ala  
 725 730 735  
 Ser Ser His Ser Tyr Pro Glu Tyr Leu His Cys Phe Asp Thr Tyr Ala  
 740 745 750  
 Pro Tyr Ile Lys Leu Asn Leu Thr Tyr Ile Arg Gln Asp Ser Phe Ser  
 755 760 765  
 Glu Lys Gly Thr Glu Gly Arg Ser Phe Asp Asp Ser Asn Leu Phe Asn  
 770 775 780  
 Leu Ser Leu Pro Ile Gly Val Lys Phe Glu Lys Phe Ser Asp Cys Asn  
 785 790 795 800  
 Asp Phe Ser Tyr Asp Leu Thr Leu Ser Tyr Val Pro Asp Leu Ile Arg  
 805 810 815  
 Asn Asp Pro Lys Cys Thr Thr Ala Leu Val Ile Ser Gly Ala Ser Trp  
 820 825 830  
 Glu Thr Tyr Ala Asn Asn Leu Ala Arg Gln Ala Leu Gln Val Arg Ala  
 835 840 845  
 Gly Ser His Tyr Ala Phe Ser Pro Met Phe Glu Val Leu Gly Gln Phe  
 850 855 860  
 Val Phe Glu Val Arg Gly Ser Ser Arg Ile Tyr Asn Val Asp Leu Gly  
 865 870 875 880

Gly Lys Phe Gln Phe  
885

<210> 14  
<211> 928  
<212> PRT  
<213> Chlamydia pneumoniae  
<400> 14

Met Lys Ser Ser Leu His Trp Phe Leu Ile Ser Ser Ser Leu Ala Leu  
1 5 10 15

Pro Leu Ser Leu Asn Phe Ser Ala Phe Ala Ala Val Val Glu Ile Asn  
20 25 30

Leu Gly Pro Thr Asn Ser Phe Ser Gly Pro Gly Thr Tyr Thr Pro Pro  
35 40 45

Ala Gln Thr Thr Asn Ala Asp Gly Thr Ile Tyr Asn Leu Thr Gly Asp  
50 55 60

Val Ser Ile Thr Asn Ala Gly Ser Pro Thr Ala Leu Thr Ala Ser Cys  
65 70 75 80

Phe Lys Glu Thr Thr Gly Asn Leu Ser Phe Gln Gly His Gly Tyr Gln  
85 90 95

Phe Leu Leu Gln Asn Ile Asp Ala Gly Ala Asn Cys Thr Phe Thr Asn  
100 105 110

Thr Ala Ala Asn Lys Leu Leu Ser Phe Ser Gly Phe Ser Tyr Leu Ser  
115 120 125

Leu Ile Gln Thr Thr Asn Ala Thr Thr Gly Thr Gly Ala Ile Lys Ser  
130 135 140

Thr Gly Ala Cys Ser Ile Gln Ser Asn Tyr Ser Cys Tyr Phe Gly Gln  
145 150 155 160

Asn Phe Ser Asn Asp Asn Gly Gly Ala Leu Gln Gly Ser Ser Ile Ser  
165 170 175

Leu Ser Leu Asn Pro Asn Leu Thr Phe Ala Lys Asn Lys Ala Thr Gln  
180 185 190

Lys Gly Gly Ala Leu Tyr Ser Thr Gly Gly Ile Thr Ile Asn Asn Thr  
195 200 205

Leu Asn Ser Ala Ser Phe Ser Glu Asn Thr Ala Ala Asn Asn Gly Gly  
210 215 220

Ala Ile Tyr Thr Glu Ala Ser Ser Phe Ile Ser Ser Asn Lys Ala Ile  
225 230 235 240

Ser Phe Ile Asn Asn Ser Val Thr Ala Thr Ser Ala Thr Gly Gly Ala  
 245 250 255  
 Ile Tyr Cys Ser Ser Thr Ser Ala Pro Lys Pro Val Leu Thr Leu Ser  
 260 265 270  
 Asp Asn Gly Glu Leu Asn Phe Ile Gly Asn Thr Ala Ile Thr Ser Gly  
 275 280 285  
 Gly Ala Ile Tyr Thr Asp Asn Leu Val Leu Ser Ser Gly Gly Pro Thr  
 290 295 300  
 Leu Phe Lys Asn Asn Ser Gly Tyr Asp Thr Ala Ala Pro Leu Gly Gly  
 305 310 315 320  
 Ala Ile Ala Ile Ala Asp Ser Gly Ser Leu Ser Leu Ser Ala Leu Gly  
 325 330 335  
 Gly Asp Ile Thr Phe Glu Gly Asn Thr Val Val Lys Gly Ala Ser Ser  
 340 345 350  
 Ser Gln Thr Thr Thr Arg Asn Ser Ile Asn Ile Gly Asn Thr Asn Ala  
 355 360 365  
 Lys Ile Val Gln Leu Arg Ala Ser Gln Gly Asn Thr Ile Tyr Phe Tyr  
 370 375 380  
 Asp Pro Ile Thr Thr Ser Ile Thr Ala Ala Leu Ser Asp Ala Leu Asn  
 385 390 395 400  
 Leu Asn Gly Pro Asp Leu Ala Gly Asn Pro Ala Tyr Gln Gly Thr Ile  
 405 410 415  
 Val Phe Ser Gly Glu Lys Leu Ser Glu Ala Glu Ala Ala Glu Ala Asp  
 420 425 430  
 Asn Leu Lys Ser Thr Ile Gln Gln Pro Leu Thr Leu Ala Gly Gly Gln  
 435 440 445  
 Leu Ser Leu Lys Ser Gly Val Thr Leu Val Ala Lys Ser Phe Ser Gln  
 450 455 460  
 Ser Pro Gly Ser Thr Leu Leu Met Asp Ala Gly Thr Thr Leu Glu Thr  
 465 470 475 480  
 Ala Asp Gly Ile Thr Ile Asn Asn Leu Val Leu Asn Val Asp Ser Leu  
 485 490 495  
 Lys Glu Thr Lys Lys Gly Thr Leu Lys Ala Thr Gln Ala Ser Gln Thr  
 500 505 510  
 Val Thr Leu Ser Gly Ser Leu Ser Leu Val Asp Pro Ser Gly Asn Val  
 515 520 525  
 Tyr Glu Asp Val Ser Trp Asn Asn Pro Gln Val Phe Ser Cys Leu Thr  
 530 535 540

Leu Thr Ala Asp Asp Pro Ala Asn Ile His Ile Thr Asp Leu Ala Ala  
 545 550 555 560  
 Asp Pro Leu Glu Lys Asn Pro Ile His Trp Gly Tyr Gln Gly Asn Trp  
 565 570 575  
 Ala Leu Ser Trp Gln Glu Asp Thr Ala Thr Lys Ser Lys Ala Ala Thr  
 580 585 590  
 Leu Thr Trp Thr Lys Thr Gly Tyr Asn Pro Asn Pro Glu Arg Arg Gly  
 595 600 605  
 Thr Leu Val Ala Asn Thr Leu Trp Gly Ser Phe Val Asp Val Arg Ser  
 610 615 620  
 Ile Gln Gln Leu Val Ala Thr Lys Val Arg Gln Ser Gln Glu Thr Arg  
 625 630 635 640  
 Gly Ile Trp Cys Glu Gly Ile Ser Asn Phe Phe His Lys Asp Ser Thr  
 645 650 655  
 Lys Ile Asn Lys Gly Phe Arg His Ile Ser Ala Gly Tyr Val Val Gly  
 660 665 670  
 Ala Thr Thr Thr Leu Ala Ser Asp Asn Leu Ile Thr Ala Ala Phe Cys  
 675 680 685  
 Gln Leu Phe Gly Lys Asp Arg Asp His Phe Ile Asn Lys Asn Arg Ala  
 690 695 700  
 Ser Ala Tyr Ala Ala Ser Leu His Leu Gln His Leu Ala Thr Leu Ser  
 705 710 715 720  
 Ser Pro Ser Leu Leu Arg Tyr Leu Pro Gly Ser Glu Ser Glu Gln Pro  
 725 730 735  
 Val Leu Phe Asp Ala Gln Ile Ser Tyr Ile Tyr Ser Lys Asn Thr Met  
 740 745 750  
 Lys Thr Tyr Tyr Thr Gln Ala Pro Lys Gly Glu Ser Ser Trp Tyr Asn  
 755 760 765  
 Asp Gly Cys Ala Leu Glu Leu Ala Ser Ser Leu Pro His Thr Ala Leu  
 770 775 780  
 Ser His Glu Gly Leu Phe His Ala Tyr Phe Pro Phe Ile Lys Val Glu  
 785 790 795 800  
 Ala Ser Tyr Ile His Gln Asp Ser Phe Lys Glu Arg Asn Thr Thr Leu  
 805 810 815  
 Val Arg Ser Phe Asp Ser Gly Asp Leu Ile Asn Val Ser Val Pro Ile  
 820 825 830  
 Gly Ile Thr Phe Glu Arg Phe Ser Arg Asn Glu Arg Ala Ser Tyr Glu  
 835 840 845

Ala Thr Val Ile Tyr Val Ala Asp Val Tyr Arg Lys Asn Pro Asp Cys  
 850 855 860

Thr Thr Ala Leu Leu Ile Asn Asn Thr Ser Trp Lys Thr Thr Gly Thr  
 865 870 875 880

Asn Leu Ser Arg Gln Ala Gly Ile Gly Arg Ala Gly Ile Phe Tyr Ala  
 885 890 895

Phe Ser Pro Asn Leu Glu Val Thr Ser Asn Leu Ser Met Glu Ile Arg  
 900 905 910

Gly Ser Ser Arg Ser Tyr Asn Ala Asp Leu Gly Gly Lys Phe Gln Phe  
 915 920 925

<210> 15  
 <211> 930  
 <212> PRT  
 <213> Chlamydia pneumoniae  
 <400> 15

Met Lys Ile Pro Leu His Lys Leu Leu Ile Ser Ser Thr Leu Val Thr  
 1 5 10 15

Pro Ile Leu Leu Ser Ile Ala Thr Tyr Gly Ala Asp Ala Ser Leu Ser  
 20 25 30

Pro Thr Asp Ser Phe Asp Gly Ala Gly Gly Ser Thr Phe Thr Pro Lys  
 35 40 45

Ser Thr Ala Asp Ala Asn Gly Thr Asn Tyr Val Leu Ser Gly Asn Val  
 50 55 60

Tyr Ile Asn Asp Ala Gly Lys Gly Thr Ala Leu Thr Gly Cys Cys Phe  
 65 70 75 80

Thr Glu Thr Thr Gly Asp Leu Thr Phe Thr Gly Lys Gly Tyr Ser Phe  
 85 90 95

Ser Phe Asn Thr Val Asp Ala Gly Ser Asn Ala Gly Ala Ala Ala Ser  
 100 105 110

Thr Thr Ala Asp Lys Ala Leu Ile Phe Thr Gly Phe Ser Asn Leu Ser  
 115 120 125

Phe Ile Ala Ala Pro Gly Thr Thr Val Ala Ser Gly Lys Ser Thr Leu  
 130 135 140

Ser Ser Ala Gly Ala Leu Asn Leu Thr Asp Asn Gly Thr Ile Leu Phe  
 145 150 155 160

Ser Gln Asn Val Ser Asn Glu Ala Asn Asn Gly Gly Ala Ile Thr  
 165 170 175

Thr Lys Thr Leu Ser Ile Ser Gly Asn Thr Ser Ser Ile Thr Phe Thr  
 180 185 190  
 Ser Asn Ser Ala Lys Lys Leu Gly Gly Ala Ile Tyr Ser Ser Ala Ala  
 195 200 205  
 Ala Ser Ile Ser Gly Asn Thr Gly Gln Leu Val Phe Met Asn Asn Lys  
 210 215 220  
 Gly Glu Thr Gly Gly Gly Ala Leu Gly Phe Glu Ala Ser Ser Ser Ile  
 225 230 235 240  
 Thr Gln Asn Ser Ser Leu Phe Phe Ser Gly Asn Thr Ala Thr Asp Ala  
 245 250 255  
 Ala Gly Lys Gly Gly Ala Ile Tyr Cys Glu Lys Thr Gly Glu Thr Pro  
 260 265 270  
 Thr Leu Thr Ile Ser Gly Asn Lys Ser Leu Thr Phe Ala Glu Asn Ser  
 275 280 285  
 Ser Val Thr Gln Gly Gly Ala Ile Cys Ala His Gly Leu Asp Leu Ser  
 290 295 300  
 Ala Ala Gly Pro Thr Leu Phe Ser Asn Asn Arg Cys Gly Asn Thr Ala  
 305 310 315 320  
 Ala Gly Lys Gly Gly Ala Ile Ala Ile Ala Asp Ser Gly Ser Leu Ser  
 325 330 335  
 Leu Ser Ala Asn Gln Gly Asp Ile Thr Phe Leu Gly Asn Thr Leu Thr  
 340 345 350  
 Ser Thr Ser Ala Pro Thr Ser Thr Arg Asn Ala Ile Tyr Leu Gly Ser  
 355 360 365  
 Ser Ala Lys Ile Thr Asn Leu Arg Ala Ala Gln Gly Gln Ser Ile Tyr  
 370 375 380  
 Phe Tyr Asp Pro Ile Ala Ser Asn Thr Thr Gly Ala Ser Asp Val Leu  
 385 390 395 400  
 Thr Ile Asn Gln Pro Asp Ser Asn Ser Pro Leu Asp Tyr Ser Gly Thr  
 405 410 415  
 Ile Val Phe Ser Gly Glu Lys Leu Ser Ala Asp Glu Ala Lys Ala Ala  
 420 425 430  
 Asp Asn Phe Thr Ser Ile Leu Lys Gln Pro Leu Ala Leu Ala Ser Gly  
 435 440 445  
 Thr Leu Ala Leu Lys Gly Asn Val Glu Leu Asp Val Asn Gly Phe Thr  
 450 455 460  
 Gln Thr Glu Gly Ser Thr Leu Leu Met Gln Pro Gly Thr Lys Leu Lys  
 465 470 475 480



Ala Asp Thr Glu Ala Ile Ser Leu Thr Lys Leu Val Val Asp Leu Ser  
 485 490 495  
 Ala Leu Glu Gly Asn Lys Ser Val Ser Ile Glu Thr Ala Gly Ala Asn  
 500 505 510  
 Lys Thr Ile Thr Leu Thr Ser Pro Leu Val Phe Gln Asp Ser Ser Gly  
 515 520 525  
 Asn Phe Tyr Glu Ser His Thr Ile Asn Gln Ala Phe Thr Gln Pro Leu  
 530 535 540  
 Val Val Phe Thr Ala Ala Thr Ala Ala Ser Asp Ile Tyr Ile Asp Ala  
 545 550 555 560  
 Leu Leu Thr Ser Pro Val Gln Thr Pro Glu Pro His Tyr Gly Tyr Gln  
 565 570 575  
 Gly His Trp Glu Ala Thr Trp Ala Asp Thr Ser Thr Ala Lys Ser Gly  
 580 585 590  
 Thr Met Thr Trp Val Thr Thr Gly Tyr Asn Pro Asn Pro Glu Arg Arg  
 595 600 605  
 Ala Ser Val Val Pro Asp Ser Leu Trp Ala Ser Phe Thr Asp Ile Arg  
 610 615 620  
 Thr Leu Gln Gln Ile Met Thr Ser Gln Ala Asn Ser Ile Tyr Gln Gln  
 625 630 635 640  
 Arg Gly Leu Trp Ala Ser Gly Thr Ala Asn Phe Phe His Lys Asp Lys  
 645 650 655  
 Ser Gly Thr Asn Gln Ala Phe Arg His Lys Ser Tyr Gly Tyr Ile Val  
 660 665 670  
 Gly Gly Ser Ala Glu Asp Phe Ser Glu Asn Ile Phe Ser Val Ala Phe  
 675 680 685  
 Cys Gln Leu Phe Gly Lys Asp Lys Asp Leu Phe Ile Val Glu Asn Thr  
 690 695 700  
 Ser His Asn Tyr Leu Ala Ser Leu Tyr Leu Gln His Arg Ala Phe Leu  
 705 710 715 720  
 Gly Gly Leu Pro Met Pro Ser Phe Gly Ser Ile Thr Asp Met Leu Lys  
 725 730 735  
 Asp Ile Pro Leu Ile Leu Asn Ala Gln Leu Ser Tyr Ser Tyr Thr Lys  
 740 745 750  
 Asn Asp Met Asp Thr Arg Tyr Thr Ser Tyr Pro Glu Ala Gln Gly Ser  
 755 760 765  
 Trp Thr Asn Asn Ser Gly Ala Leu Glu Leu Gly Gly Ser Leu Ala Leu  
 770 775 780

Tyr Leu Pro Lys Glu Ala Pro Phe Phe Gln Gly Tyr Phe Pro Phe Leu  
 785 790 795 800  
 Lys Phe Gln Ala Val Tyr Ser Arg Gln Gln Asn Phe Lys Glu Ser Gly  
 805 810 815  
 Ala Glu Ala Arg Ala Phe Asp Asp Gly Asp Leu Val Asn Cys Ser Ile  
 820 825 830  
 Pro Val Gly Ile Arg Leu Glu Lys Ile Ser Glu Asp Glu Lys Asn Asn  
 835 840 845  
 Phe Glu Ile Ser Leu Ala Tyr Ile Gly Asp Val Tyr Arg Lys Asn Pro  
 850 855 860  
 Arg Ser Arg Thr Ser Leu Met Val Ser Gly Ala Ser Trp Thr Ser Leu  
 865 870 875 880  
 Cys Lys Asn Leu Ala Arg Gln Ala Phe Leu Ala Ser Ala Gly Ser His  
 885 890 895  
 Leu Thr Leu Ser Pro His Val Glu Leu Ser Gly Glu Ala Ala Tyr Glu  
 900 905 910  
 Leu Arg Gly Ser Ala His Ile Tyr Asn Val Asp Cys Gly Leu Arg Tyr  
 915 920 925  
 Ser Phe  
 930

<210> 16  
 <211> 293  
 <212> PRT  
 <213> Chlamydia pneumoniae  
 <400> 16

Met Leu Ser Ser Leu Ile Arg Asp Ser Phe Pro Leu Leu Ile Leu Leu  
 1 5 10 15  
 Pro Thr Phe Leu Ala Ala Leu Gly Ala Ser Val Ala Gly Gly Val Met  
 20 25 30  
 Gly Thr Tyr Ile Val Val Lys Arg Ile Val Ser Ile Ser Gly Ser Ile  
 35 40 45  
 Ser His Ala Ile Leu Gly Gly Ile Gly Leu Thr Leu Trp Ile Gln Tyr  
 50 55 60  
 Lys Leu His Leu Ser Phe Phe Pro Met Tyr Gly Ala Ile Val Gly Ala  
 65 70 75 80  
 Ile Phe Leu Ala Leu Cys Ile Gly Lys Ile His Leu Lys Tyr Gln Glu  
 85 90 95

Arg Glu Asp Ser Leu Ile Ala Met Ile Trp Ser Val Gly Met Ala Ile  
 100 105 110  
 Gly Ile Ile Phe Ile Ser Arg Leu Pro Thr Phe Asn Gly Glu Leu Ile  
 115 120 125  
 Asn Phe Leu Phe Gly Asn Ile Leu Trp Val Thr Pro Ser Asp Leu Tyr  
 130 135 140  
 Ser Leu Gly Ile Phe Asp Leu Leu Val Leu Gly Ile Val Val Leu Cys  
 145 150 155 160  
 His Thr Arg Phe Leu Ala Leu Cys Phe Asp Glu Arg Tyr Thr Ala Leu  
 165 170 175  
 Asn His Cys Ser Val Gln Leu Trp Tyr Phe Leu Leu Leu Val Leu Thr  
 180 185 190  
 Ala Ile Thr Ile Val Met Leu Ile Tyr Val Met Gly Thr Ile Leu Met  
 195 200 205  
 Leu Ser Met Leu Val Leu Pro Val Ala Ile Ala Cys Arg Phe Ser Tyr  
 210 215 220  
 Lys Met Thr Arg Ile Met Phe Ile Ser Val Leu Leu Asn Ile Leu Cys  
 225 230 235 240  
 Ser Phe Ser Gly Ile Cys Ile Ala Tyr Cys Leu Asp Phe Pro Val Gly  
 245 250 255  
 Pro Thr Ile Ser Leu Leu Met Gly Leu Gly Tyr Thr Ala Ser Leu Cys  
 260 265 270  
 Val Lys Lys Arg Tyr Asn Pro Ser Thr Pro Ser Pro Val Ser Pro Glu  
 275 280 285  
 Ile Asn Thr Asn Val  
 290